

Harald H. Zimmermann / Volker Schramm (Hrsg.)

Knowledge Management und Kommunikationssysteme

**Proceedings des 6. Internationalen Symposiums
für Informationswissenschaft (ISI '98)
Prag, 3.-7. November 1998**

**Hochschulverband für Informationswissenschaft (HI) e.V. Konstanz;
Fachrichtung Informationswissenschaft der Universität des
Saarlandes, Saarbrücken**

UVK Universitätsverlag Konstanz

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Knowledge Management und Kommunikationssysteme :

Workflow Management, Multimedia, Knowledge-Transfer ;
proceedings des 6. Internationalen Symposiums für
Informationswissenschaft (ISI '98) Prag, 3. – 7. November 1998 / <
Hochschulverband für Informationswissenschaft (HI) e.V. Konstanz ;
Fachrichtung Informationswissenschaft der Universität des Saarlandes, Saarbrücken.
Harald H. Zimmermann / Volker Schramm (Hg.).
Konstanz: UVK, Univ.-Verl. Konstanz, 1998

Schriften zur Informationswissenschaft; Bd. 34)

ISBN 3-87940-653-7

ISSN 0938-8710

ISBN 3-87940-653-7

© UVK Universitätsverlag Konstanz GmbH, Konstanz 1998

Satz: Reproduktionsfertige Vorlagen von den Herausgebern

Druck und Bindung: Digital Druck GmbH, Frensdorf

Einbandgestaltung: Riester & Sieber, Konstanz

Inhalt

ISI '98	9
Section 1 Workflow Management (Modeling of Workflow, Decision Support, Knowledge Management)	
<i>Emilia Currás</i> Dialectics in Knowledge Organization: An Approach	17
<i>Mohsen Shafaei</i> „KHSFlow“ Ein System zur hypertextbasierten Workflowmodellierung und –steuerung	32
<i>Rainer Hammwöhner</i> Entscheidungsunterstützung bei der Planung von Übersetzungsprojekten	47
<i>Leo Vodáček</i> Information Management: Concept, Teaching, Applications	58
<i>Elisabeth Davenport, Kathy Buckner</i> SO-grams: a personal visualisation toolkit for intranet users	67
<i>Patric Hoechst, Georg Schneider, Karsten Merschjan, Jean Schweitzer</i> "Telearbeit auf der Grundlage einer neuen Kommunikationstechnologie"	80
<i>Otto Krickl, Elisabeth Milchrahm</i> Wissensmanagement in Lernenden Organisationen Eine neue Herausforderung für die Informationswissenschaft?	88
Section 2 Multimedia and Knowledge Transfer (Electronic Market, Knowledge Exchange)	
<i>Josef Herget, Silke Grossmann</i> Bedarf, Architektur und Nutzen elektronischer Marktplätze - einige Konstruktionshinweise dargestellt am Beispiel des Marktplatzes Bildung	105
<i>Sylva Simsova</i> Between the lines: Newsgroups as an information source for social historians	123
<i>Gunter Bauer, Gerhard Reichmann</i> Projekt "Infopoint"- Errichtung eines Informationskiosks an der Universität Graz	137
<i>Jela Steinerová</i> Interpreting Knowledge in Information Science	143

<i>Gerhard Knorz</i> Dynamische Menüs zur Recherche und Orientierung. Konzeption der thematischen Suche für den Informationsservice "WebSite 'Methodik'"	154
<i>Claus Rautenstrauch, Jochen A. Siegle</i> Online-Marketing von Rundfunkmedien – Ergebnisse einer empirischen Studie	168
<i>Werner Schweibenz</i> The "Virtual Museum": New Perspectives For Museums to Present Objects and Information Using the Internet as a Knowledge Base and Communication System	185
<i>Andreas Dietzsch, Werner Esswein</i> Wissensaustausch auf der Grundlage von Referenzmodellen – Anforderungen und Möglichkeiten	201
Section 3 Electronic Publishing (Impact on Libraries, Structured Documentation)	
<i>Stefan Kimlicka</i> Bibliographic database as a support tool for knowledge management	213
<i>T.D. Wilson</i> Information Research: a case study in the free electronic publication of research	225
<i>Jiri Panyr, Werner Zucker</i> Informations- und Wissensmanagement durch strukturierte Dokumente	235
<i>Saša Skenderija</i> Library within Cyberspace Paradigm: the Integrated Virtual Library Concept	253
<i>Vladimir T. Borovansky</i> Support of the Information Needs of Scientists and the changing Pattern of scholarly Communication	256
<i>Karel Sosna</i> Electronic Library "Czech Parliament"	258

Section 4 The Internet as Knowledge Base (Social Aspects and Impacts, Interface Problems)

<i>Wolfgang Semar</i> Eine empirische Studie über die Auswirkungen elektronischer Marktplätze	263
<i>A.L. Heuer, K. Siemonsen, T. Engel and C. Meinel</i> Interface for Navigation and Database - Querying via WWW	273
<i>Rainer Kuhlen, Anja Odenthal</i> Elektronische Kommunikationsforen – das Beispiel UNESCO VF-INFOethics	283
<i>Alexander Sigel</i> Long-Term Value Adding in an Open Category Network: An Informal Social Approach Towards Relating Conceptual Order Systems on the Internet	296
<i>M. Burke</i> Teddy Bears and Perfumes: Advertising Information - Knowledge Management at its' best!	305
<i>Thomas Mandl</i> Vague Transformations in Information Retrieval	312

Section 5 Recent Developments of Information Systems and Tools (Quality Aspects)

<i>Alexander Kaiser</i> Neuere Entwicklungen auf dem Gebiet temporaler Datenbanken - eine kritische Analyse	329
<i>Fahri Yetim</i> Cultural Aspects of Group Support Systems	344
<i>Jutta Marx, Bernhard Staudinger</i> Das Konzept einer strikt benutzerorientierten Evaluierung maschineller Übersetzungssysteme (am Beispiel des Projekts MIROSLAV)	358
<i>Christian Wolff</i> Elektronische Wahlverfahren als Beispiel angewandter Kryptographie	373
<i>Rolf Assfalg</i> Rhetorische Elemente in hypertextbasierten elektronischen Foren	385
<i>Hubert Hüther</i> Selix im DFG-Projekt Cascade	397

<i>Ruppert von Teutul, Georg Schneider, Jean Schweitzer</i> Team Finder: Ein intelligenter Assistent zur zielgerichteten Ressourcen- selektion für workflow-integrierte multimediale Audio- / Video-Desktop- konferenzen	404
<i>Jürgen Krause, Thomas Mandl, André Schaefer, Maximilian Stempfhuber</i> Text-Fakten-Integration in Informationssystemen	413
<i>Osman Bayraktar, Christa Womser-Hacker</i> Qualitätsbewertung von Information-Retrieval-Systemen Sind Synergieeffekte durch die Koordination verschiedener Bewertungsmethoden möglich?	427
<i>Bernard Bekavac, Marc Rittberger</i> Ein Navigationsassistent für das World Wide Web	438
Section 6 The Information Society (Different Aspects)	
<i>Frieder Nake</i> Schwierigkeiten beim semiotischen Blick auf die Informationsgesellschaft	455
<i>Antonín Rosický</i> Information Problems in Information Society	469
<i>Jiří Cejpek</i> Informationsgesellschaft oder eine andere Gesellschaft?	481
<i>Elena Macevièiūtė</i> Manpower Requirements in the Emerging Information Society of the Baltic States	486
<i>Rudolf Vlasak, Stanislav Kalkus</i> Information Sector as Economical Category	501
<i>Gerhard Reichmann</i> Schutz von Daten und Schutz von Wissen in der Informationsgesellschaft	508
<i>Hans W. Giessen</i> Telekommunikation und Partizipation	523
<i>Gerhard Fröhlich</i> Optimale Informationsvorenthaltung als Strategem wissenschaftlicher Kommunikation	535
Autoren-Index	551



6.

Internationales
Symposium für
Informationswissenschaft

3.-7. November 1998
Karls-Universität zu Prag

Knowledge Management und Kommunikationssysteme

Veranstalter

Hochschulverband für Informationswissenschaft (HI) e.V. Konstanz
Fachrichtung Informationswissenschaft der Universität des Saarlandes,
Saarbrücken.

Programmkomitee

Prof. Dr. Harald Zimmermann, Uni
Saarbrücken (Vorsitz)
Dr. Rolf Assfalg, Uni Konstanz
Prof. Dr. Raffael Capurro, FH Stuttgart
Prof. Dr. Jiri Cejpek, Karls-Universität
Prag
Prof. Dr. Hans-Peter Frei, UBS Zürich
Prof. Dr. Norbert Fuhr, Uni Dortmund
Dr. Hans Giessen, Uni Saarbrücken
Prof. Dr. Rainer Hammwöhner, Uni
Regensburg
Dr. Ilse Harms, Uni Saarbrücken
Prof. Dr. Ralf-Dirk Hennings, FH
Potsdam
Prof. Dr. Norbert Henrichs, Uni
Düsseldorf
Dipl. Soz. Matthias Herfurth, IZ Bonn
Dr. Josef Herget, EMS Konstanz
Stephan Holländer, HTL Chur
Dr. Stanislav Kalkus, Karls-Universität
Prag
Prof. Dr. Gerhard Knorz, FH Darmstadt
Prof. Dr. Alfred Kobsa, GMD FIT
St. Augustin
Prof. Dr. Jürgen Krause, IZ Bonn

Prof. Dr. Rainer Kühlen, Uni Konstanz
Dr. Heinz-Dirk Luckhardt, Uni
Saarbrücken
Prof. Dr. Achim Oßwald, FH Köln
Dr. Jiri Panyr, Siemens AG München
Peter Poschadel, Uni Saarbrücken
(stud. Vertreter)
Prof. Dr. Wolf Rauch, Uni Graz
PD Dr. Ulrich Reimer, Swiss Life
Zürich
Prof. Dr. Harald Reiterer, Uni Konstanz
Dr. Marc Rittberger, Uni Konstanz
Dr. Christian Schlögl, FH Eisenstadt
Prof. Dr. Ralf Schmidt, FH Hamburg
Prof. Dr. Eric Schoop, TU Dresden
Prof. Dr. Thomas Seeger, FH
Darmstadt
Thomas Tanzer, ETH Lausanne
Dr. Stephanie Teufel, Uni Zürich
Dr. Ulrich Thiel, GMD IPSI, Darmstadt
PD Dr. Rudolf Vlasak, Karls-
Universität Prag
Prof. Dr. Gernot Wersig, FU Berlin
PD Dr. Christa Womser-Hacker, Uni
Konstanz

Organisationskomitee

Karls-Universität, Prag,

Universität des Saarlandes, Saarbrücken

Prof. Dr. Jiri Cejpek

Dr. Stanislav Kalkus,

PD Dr. Rudolf Vlasak

Prof. Dr. Harald Zimmermann

Dr. Volker Schramm

Dank!

Bei allen Helferinnen und Helfern, die mit zum Gelingen dieses Tagungsbandes beigetragen haben, bedanke ich mich herzlich. Es ist mir jedoch ein Bedürfnis, die Damen und Herren Dr. Rolf Assfalg, Kornelius Bamberger, Walter Heinrich, Patric Hoechst, Monika Lotzen, Dr. Heinz-Dirk Luckhardt, Ursula Markus, Stefania Racioppa, Dr. Matthias Schüler, Martin Schulze, Astrid Tessmer und Ruppert von Teutul für ihren unermüdlichen Einsatz ganz besonders hervorzuheben.

Volker Schramm



650 anni

UNIVERSITATIS CAROLINAE

1348-1998

Prof. JUDr. Karel Malý, DrSc.
Chancellor of the Charles University, Prague

Ladies and gentlemen,

may I welcome you on the grounds of the venerable Carolinum, the seat of the office of the Charles university Chancellor. Charles the Fourth, the Czech king, established this university 650 years ago in the year 1348. Certain symbolism can be observed in the fact that the oldest and largest library of the university - the library of the Charles college - was transferred in 1983 to the house of a wealthy Prague mint master Rotlev. His house was standing partially on the grounds of this aula. European university libraries of the Middle Ages, respectively the college libraries, represented then a new era in creating conditions for development of not only Czech, but also European education. After all the European universities of Middle Ages were truly international institutions.

I assume that your Sixth International Symposium on Information Sciences with a view toward management of knowledge and of communication systems, wants to create something similar, of course in terms of the end of 20th Century: to create conditions toward achievement of higher public and special education. Understandably by other means than the manuscript of Middle Ages. Furthermore creation of information communication hinterland not only in academic environment but in the whole organism of democratic states worldwide, is at stake.

Conflict of two developing orders is the most outstanding feature of the end of the 20th Century, the order of the Earth, and the order of the human culture in the wide framework of all that was created by man what affected his environment. It is not the biosphere that is endangered by this conflict, it was here billions of years before man but by man created culture with all the time increasing antinatural elements.

You as moderators of information communication processes, collecting, processing, and disseminating recorded knowledge of past and present generations are able to moderate this conflict. For this task you have tools today of which your predecessors could not even dream about.

I wish you that here in these surroundings which remind us of the best traditions of European education, your meeting will contribute toward enlightenment and solution of actual global problems, in this case by taking advantage of achievements of the modern information science.

ISI'98 IN THE CZECH ACADEMIC SURROUNDING - ÚISK AS HOST OF THE CONFERENCE

Doc. PhDr. Rudolf Vlasák

ÚISK Director

In Prague, the capital city of the Czech Republic, and especially on the grounds of its Charles university it is customary to remind of outstanding and deeply historical traditions. The graduate education in the information fields, although deriving from the classical librarianship, cannot rightfully be included in this, it is nevertheless one of the most modern disciplines but it still has relatively solid tradition.

First library courses in the independent Czechoslovak Republic were developed since 1927. Two state library schools were established and lectures in librarianship were started at the Charles University, Faculty of Philosophy. After closing of universities during German occupation, in years 1945-1948 this tradition was continued by partial renewal of library education. After establishment of totalitarian regime in 1948 and under the influence of Soviet library education, structural changes were introduced and so was the ideology of education in the spirit of state ideology of the ruling communist party.

Already in December 1989, one month after the "velvet revolution" in which university students played an extremely significant role a commission was established to institute changes of educational program. Within four months it did prepare new education program in such a way so it would be possible to project it as soon as possible into teaching. The main changes consisted in removing existing ideological limits and opening the space for tolerance in the free dissemination of knowledge in interests of development of spiritual, cultural, scientific, technical and economical renewal of society. The Institute of Information and Library Studies (ÚISK) was established.

The Institute is at present the principal Czech graduate institution with full program of field studies (Bachelor, Master and PhD) in its full range - from library science and historical bibliography, through public librarianship to information science and information management. It is organically incorporated into the overall education system of the Czech Republic as a standard educational and scientific research institute at the very traditional faculty of the biggest and most significant university in the country. We are engaged in the research tasks in the frame of national and also international research and development programs, participate on the exchange contracts of teachers and students with the foreign universities, fulfilled the role of the Czech National Member of FID, each year organise the conference with international participation "Library Automation" and our students have last year started to publish the first Czech full electronic professional journal named "Ikaros".

The nature of work of all information institutions is influenced by new information technologies and most of all by democratization of access to electronic information resources. The space for information specialist or librarian as intermediary between user and information source is constantly

narrowing. On the other hand the need to organize systems of creation, maintenance, and development of interfaces for access to these programs ought to be adjusted to these needs.

Present educational concept at ÚISK stems from basic and mutually balanced specialized division of education after common two year basic program directed toward library science, librarianship and information studies. Here the students are systematically prepared for a very wide spectrum of employment which we believe has a future in the next century.

In the subjects aimed to planning and design of information systems we teach that we define the system processing operations after the definition of output. Also the structures and contents of education and training of new professionals we derive from the concept of the profession.

These can be summarized in four wide groups:

1. Literary experts knowledgeable of the early history of manuscripts up to the present forms of the books and other documents including subjects related to their protection, restoration, further library processing and evaluation.
2. Librarians in public libraries, who are also active in a given locality as cultural representatives participating at organizing free time programs for citizens of all age categories.
3. Managers and leading specialists in special libraries, research and academic libraries or information centers serving via all contemporary modern forms of access to scientific and technical information, but also in graduate education in library and information science.
4. Information managers active practically in all types and kinds of enterprises and institutions as organizers of creation, access, selection and use of information sources, eventually active as independent entrepreneurs (information consultants).

It concerns evaluation of the labour market in Czech Republic and so far it shows that in the above mentioned fields it is still open in spite of increasing percentage of unemployment. The fact that the announcement table placed in administration spaces in our institute offering employment in our fields is all the time full of listed vacancies, is witness to that. Most of the students find the future employment already during their studies and often start to work on part time bases. Those interested in more detailed information about our institute I would refer to our homepage URL: <http://www.cuni.cz/~skenders/uisk>.

This year ISI honored ÚISK and Charles University, which is celebrating 650 years of its existence, in letting it to organize the conference. We are the first non German speaking country, which was given this opportunity. We appreciate this and sincere welcome all its participants. We believe that the atmosphere of this historically famous university and also the autumn Prague will create plesurable framework for professional and also cultural impressions.

Vorwort

Erstmals findet das 'International Symposium on Information Science' (ISI) außerhalb des deutschsprachigen Raumes statt; damit wird der internationale Charakter auch äußerlich betont. Wiederum kommen viele Beiträge - wie bisher schon - aus Deutschland, manche werden auch in deutscher Sprache gehalten. Ein Grund dafür, an Deutsch als *einer* der Kommunikationssprachen (neben Englisch) festzuhalten, ist die einfachere Zugänglichkeit der Überlegungen und Ergebnisse im deutschsprachigen Raum.

Wissens-Management (Knowledge Management) - das Leitthema der Tagung - erscheint manchem vielleicht als moderner Terminus oder auch neuer Gegenstandsbereich. Im informationswissenschaftlichen Umfeld sind sowohl die grundlegende Bedeutung als auch Problematik jedoch seit langem vertraut; ausgeweitet haben sich in erster Linie die informationstechnischen Möglichkeiten und damit verbunden auch die Einsatzfelder, wobei die Integration in betriebliche Prozesse einen besonderen Schwerpunkt darstellt. Mit A.W. Scheer konnte als Gastredner zum Hauptthema ein anerkannter Wissenschaftler der Wirtschaftsinformatik gewonnen werden, der Theorie und Praxis erfolgreich zu verbinden vermag.

Es war nicht immer leicht, die Beiträge den einzelnen vorgesehenen Sektionen zuzuordnen; zudem waren auch die organisatorischen Gegebenheiten in die Planung einzubringen. Ein Symposium soll die wissenschaftliche Diskussion in interessanten Themenfeldern anregen, die Beiträge stellen also keine Anthologie der jeweiligen Themenfelder dar, sondern eher Exempla zum State-of-the-Art. Andererseits ist in diesem Themenfeld kaum eine andere Tagung zu finden, bei der die *ganzheitliche* Betrachtung (wirtschaftliche, technische, soziale ... Aspekte der Informationsgesellschaft) so sehr im Vordergrund steht. Der Tagungsband gibt daher jedem - so steht zu hoffen -, der einmal über den Tellerrand schauen möchte, eine Reihe von Anregungen.

Es ist mir eine liebe Pflicht, an dieser Stelle meinem Mitarbeiter, Dr. Volker Schramm, für seinen unermüdlichen Einsatz bei der Organisation zu danken. Ohne ihn hätte meine vor Jahren gemachte Zusage, die ISI '98 in Prag mit zu betreuen, nicht sachgerecht realisiert werden können.

Saarbrücken, den 21.9.1998

Harald H. Zimmermann

Section 1

Workflow Management

(Modeling of Workflow, Decision Support, Knowledge Management)

DIALECTICS IN KNOWLEDGE ORGANIZATION: AN APPROACH

Emilia Currás

C/ O'Donnell 6, C-12, 28009-Madrid. Spain.

SUMMARY

This is a study of dialectics in its own right. Originally, it was a pedagogical means of teaching people to think, reason and engage in dialogue. Later on it became a part of logic and at present it is applied to any human activity, whether intellectual or social, to put forward and contrast arguments which clarify the situations in question.

Mention is made to the processes of thesis, antithesis and synthesis by means of which something which can be, or has already been demonstrated, is then argued, analysed, seeing the different possible points of concordance or discordance in order to reach a conclusion.

This method is applicable, in particular, to the organisation of knowledge in information sciences, such as drafting summaries, indexing, content analysis, classification processes, among other information science techniques, where a text is analysed, minutely examined, pointing out the pros and cons, with an end result in view, whether this be a summary, a few key words which determine the content for database construction, or, for classification purposes, etc.. Finally, it is deduced that there is a case for an approach to "dialectics in information science".

Key words: dialectics, dialectics in information science, dialectics in knowledge organization, information systems, information science summaries, content analysis, classification processes, data base construction.

The study of dialectics, by means of its features and applications, can prove to be very interesting, almost exciting. Although over the last decades of the twentieth century it has come back into vogue and applications can be found in all branches of knowledge and even in everyday life, dialectics is as old as Humanity itself. We could say that men thought in a dialectic way before they were aware of doing so and before its etymology, definitions and applications were studied. The same way that they used information without knowing what it was, nor which connotations or practices could formally be attributed to it. Nevertheless, in our Western civilisation, we know that around

the time of the Greeks dialectics and its applications began to be studied and this continued, on and off, throughout History. It worth noting that the periods of greater interest have coincided with those periods in history of accelerated evolution, of transmutations in the manifestations of thought or of socio-political crises. Nowadays, humankind is going through all these processes at the same time and precisely that is why dialectics has once again come to form a part of the theories and practices of our historical evolution.

That is also why this paper deals with and applies dialectics to the organisation of knowledge in information science version. In previous papers, dialectics has been applied to several subjects, such as, the interaction of sciences (Currás, 1990-1996), the relationship between information - information science (Currás, 1990-1995), as well as the study of systems science (Currás 1996, 1993) amongst other works.

ETYMOLOGY

The term "dialectics" comes from the Greek word "dialogue" which is a compound of (Fig. 1)


dia = between

and

legein = to converse

The term "dialectic" (Ferrater, 1981) is made up of

- dialectique (hispanicism of the Greek word)

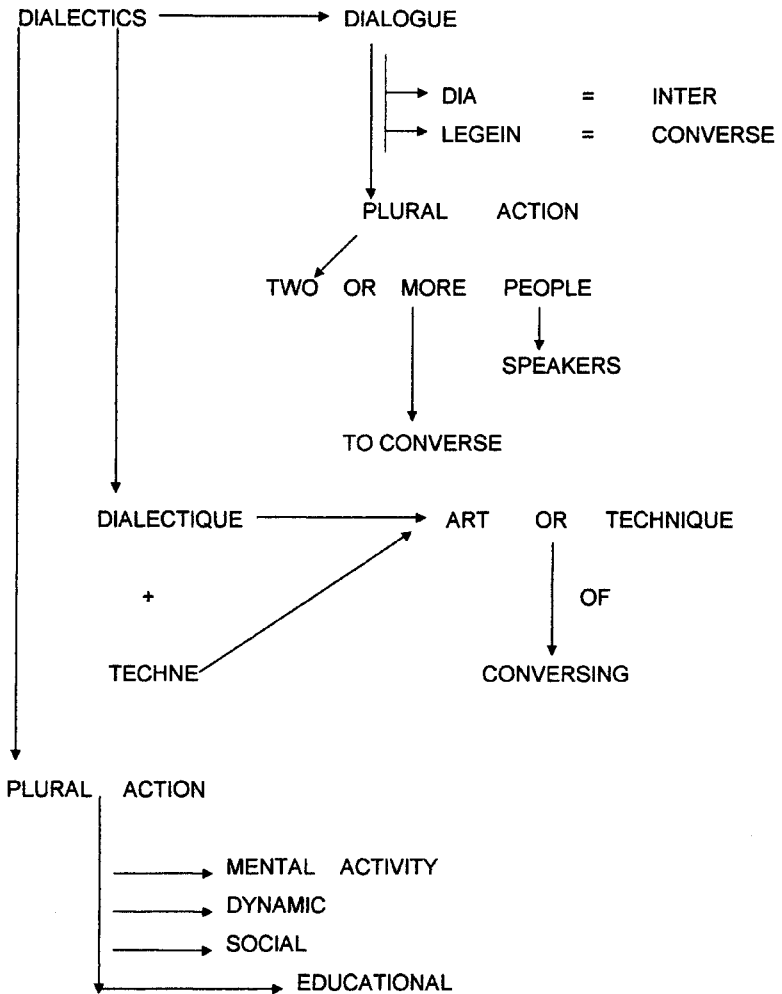
which means, to be precise ,  dialogue

and

- techne (again a hispanicism)

which means  technique or art

ETYMOLOGY



In other words, we deal with the technique or art of conversing, which throughout History has undergone a series of changes in its theoretical concept.

DEFINITIONS

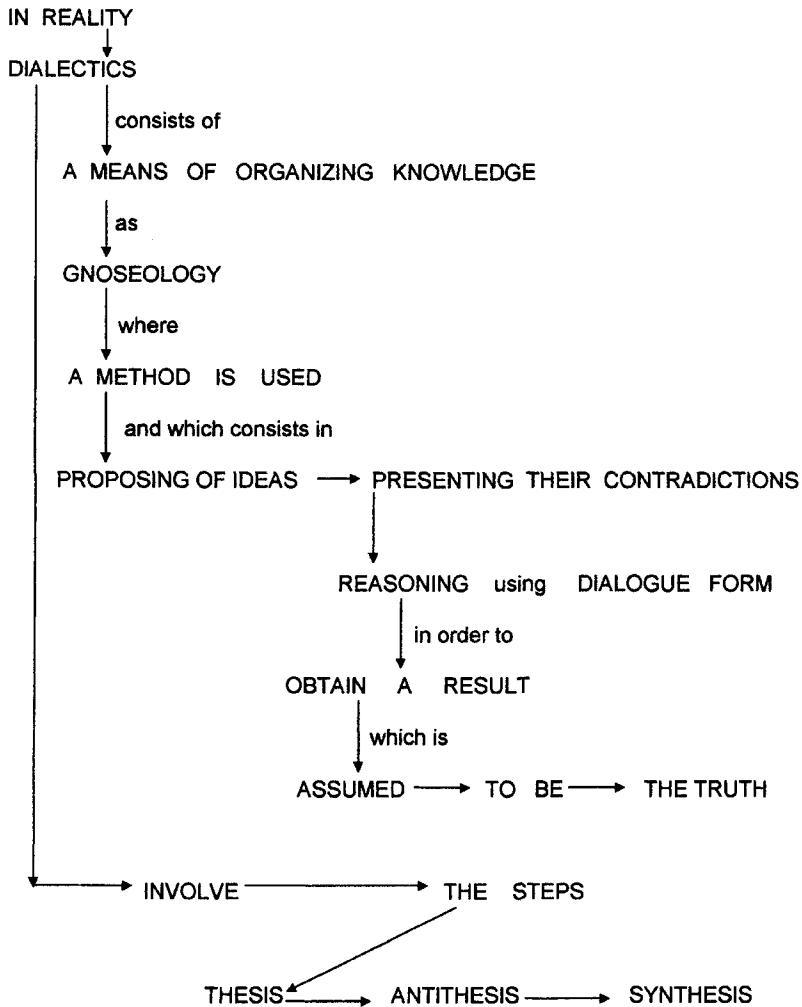
The vast amount of definitions on the subject given in the bibliography can be reduced to a few since certain principles have prevailed through time. We will point out the more classic ones and others which are more closely related to information science in the organization of knowledge.

Dialectics (definitions, given in chronological order).

- what is dialectics if not skill and expertise in arguing (St. Augustine, 354-430, Núñez Ruiz, 1975).
- it is the study of the theory of knowledge, which deals with the cognoscitive relationship between object and subject (Livio Sichirollo, 1976).
- the art of demonstrating a proposed thesis by means of the classification of concepts and the rigorous distinction between them (Diego Núñez Ruiz, *1975).

* It is very difficult to find biographical information on contemporary and recently deceased authors.

DEFINITIONS



- the art of speaking, reasoning, discussing, conversing, putting forward controversial ideas. It is a purely mental activity of the human brain in its discursive, not intuitive, mode (Capone Braga, 1957).

As can be seen, by extrapolating, we can find a relationship between dialectics and information science in its organisation of knowledge, e.g. content analysis, summaries, classification, etc.

In reality, dialectics is, in general, a means of organizing knowledge like a "gnoseology" which consists of proposing ideas by putting forward their opposite, reasoning - using them in dialogue -, to obtain a result which is assumed to be the truth one. We can understand, from this, the steps set out by Plato (427-347 B.C., Ferrater, 1981) and Hegel (1770-1831, Sandor, 1964) amongst others, and which are:

thesis

antithesis

synthesis

These processes are easily applied to organization of knowledge. (Fig. 2)

DIALECTICS RECENT TRENDS

In the last decades of the 20th century Mankind has passed through an historical period of great change, moving from the particular to the general and viceversa. Thinking has been globalised and raised to a higher degree of abstraction, which allows us to see the whole in which each particular element evolves. We can classify this present moment in which we live as highly dialectic. As a result, the sphere of dialectics, in its conception and applications has widened. One way or another, new branches of knowledge have sprung up, using dialectics as part of their evolution and discourse. For the purposes of this paper, we are interested in the empirical or scientific dialectics and the systems science point of view.

- Scientific dialectics was developed by Ferdinand Gonseth (1890-, Rodríguez Delgado, 1997), the Swiss mathematician and philosopher, editor of the magazine "Dialectics" and founder of a School of thought. His originality consisted in setting down the bases for dialectics on the principles of complementariness and not contradiction. To a certain extent, he followed the ideas of Pythagoras (582 B.C., *Enciclopedia filosófica*, 1996) and Hegel (1770-1831, Sandor, 1964). He tried to overcome the contradictions derived from dialectic idealism and materialism by means of a synthesis which encompassed and reduced opposites, considering them as extreme aspects of one continuous and as complementary aspects of a reality seen from various angles. It is applied to intuitive-rational knowledge, object-subject, individual-society...

The Gonseth School tries to embrace all branches of knowledge and dissolve the dichotomies to obtain a global view (perception) which integrates (dialectic integration) all the opposites and contradictions found in science and in society.

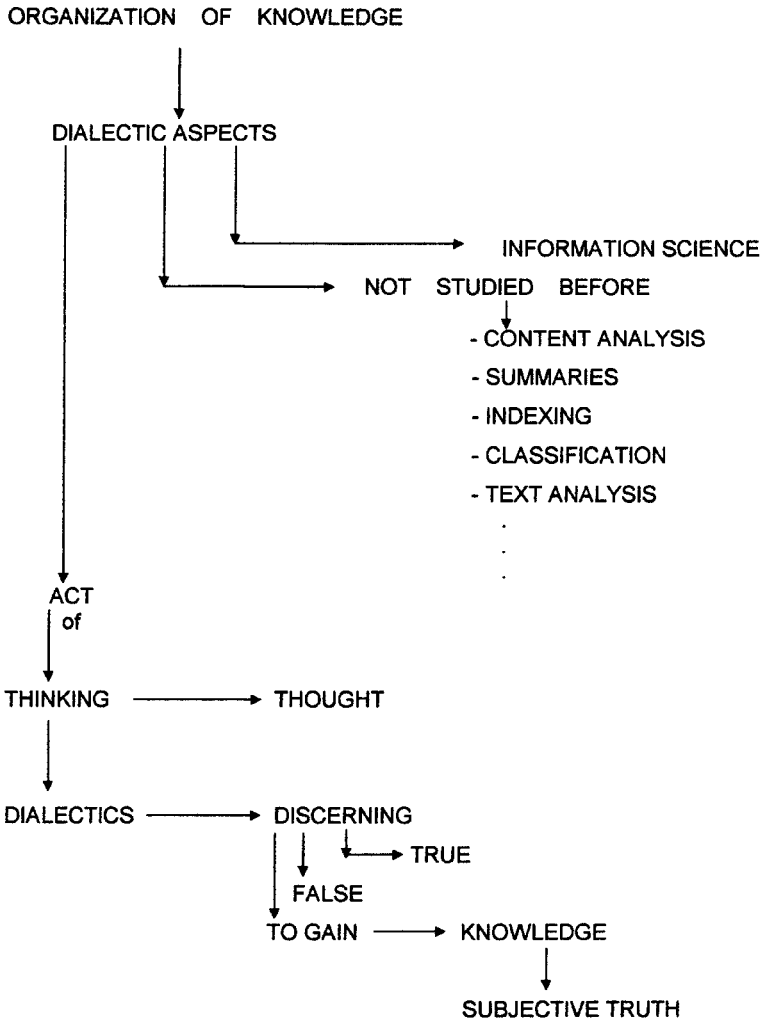
- Systems dialectics: In this day of globalisation, the world is seen as a system in which the constituent elements are interrelated, and evolve according to interactions in time and place, both within and outside the system.. Rafael Rodríguez Delgado who introduced the present Theory of Systems (systems science) and its detailed study into Spain (1980), (Rodríguez Delgado, 1997),

observed that not all evolutionary systems followed a continuous pattern of development. They could not be studied by applying the laws of systems dynamics. These are systems which undergo transformations, which change their development by producing contradictory evolutions, which are also at times concordant and discordant. They, therefore, fulfill the conditions for being considered as dialectic.

Rodríguez Delgado introduces a new variable "the transformable" and postulates systems dialectics which can be applied to all branches of science. In previous papers (Currás, 1990-1996) I have spoken about systems dialectics applied to information science and various other questions.

DIALECTIC ASPECTS IN THE ORGANIZATION OF KNOWLEDGE

Bear in mind that, in general, we are faced with a question which had not been applied until now to knowledge organisation in information science ; that is, knowledge organisation per se, as applied to document analysis, to summary and classification techniques, under other information science techniques, it would, therefore, be advisable to lay down the principles, definitions and historical changes on which the assertions which follow are based. (Fig. 3)



DIALECTICS IN THE ORGANIZATION OF KNOWLEDGE

FIG. 3

The act of thinking and thought is linked to dialectics insofar that the latter is a means of discerning between true and false - true thought, false thought, the wrong hypothesis - to understand things, the subjective, human truth. The first pages of this paper include some definitions of dialectics, linking it directly to ways of thinking. Let me quote as a good example, the second of the definitions listed: "it is the study of the theory of knowledge, which deals with the cognoscitive relationship between object and subject" (Livio Sichirollo, 1976).

As for the theory of knowledge, Kant (1724-1804, Sánchez Ortiz, 1977) tells us that knowledge begins with the senses and ends with reason. This can bear relation to the creation of knowledge by the mind, by the influence of information (quanta of information) reaching it, beginning with the thought process, to produce a form of subjective knowledge. This relationship between information, reasoning, thought, knowledge can be extrapolated from some of Aristotles (384-322 B.C., Sánchez Ortiz, 1977) ideas. For example, when he deals with the origin of knowledge, based on sensitive perception by means of abstracting understanding to reach universal ideas.

The theory and organisation of knowledge are derived from thought as a means of studying the cognoscitive relationship between subject and object. In the particular case of knowledge organisation in relation to information science it must be understood that the subject is the starting point and the object the end or result. The former could quite easily be the information expert, the analyst, the indexer, classifier, etc. or a concrete text. The object will be the proposed end product, such as: a summary, a catalogue, a database, as examples between other techniques.

DIALECTIC METHODOLOGY

The methodology of dialectics derives from certain processes, which starts with something concrete, the *essence*, which is or exists, then passes thorough experience, concordant and discordant lines of argument, while leading, at the same time, to *conscience*, getting the point of a *final postulate*. (Fig. 4)

These steps are parallel to the classic steps put forward by Plato (*Enciclopedia filosófica*, 1996, Sandor, 1964) and Hegel (Engels, 1968, Sandor, 1964), among others, and which establish a *thesis*, *antithesis*, *synthesis*. Hegel (*Enciclopedia filosófica*, 1996, Sandor, 1964) explains these concepts more clearly when he refers to the *theory of being* (original object), the *theory of essence* (that which is or is not - analysis) and *theory of concept* (what it becomes - point of resolution). (Fig. 4)

DIALECTICS IN INFORMATION SCIENCE: DIALECTIC CORRELATIONS IN THE ORGANISATION OF KNOWLEDGE TECHNIQUES

After reading the above on will deduce that both dialectics, and its methodology can be applied perfectly well to the processes of content analysis, summary confection, classification, and indexation techniques, cataloguing..., among other information science techniques.

All of them can be, and in fact are, carried out, following the steps given above. For example, an information science analyst who wants to draw up a summary to build up a database, starts with a given texts - *thesis*, subject, *what it is* -.

He analyse the texts, he breaks them down into parts - *antithesis* - negation, discordance, experimentation, *what could be* -.

And then, in a process of *synthesis* put the parts back together again - concordance, resolution, *what it becomes* - and draws up his summary - *the object*. All of this is a dialectic process. In other words, dialectics enables us - in a process of division - analysis of a set text - to compose - synthesis - another text of different characteristics (Pinto, 1992, 1991).

The same processes are followed when we want to classify a text, let's say, a book. We take a book, we analyse it in different stages, considering the prologue, chapters thema, etc. and afterwards the pieces are recomposed to decide how to classify it.

In the case of classification techniques, the dialectic process is a very complex one, since it is not easy to see in which category to put a book or any other document. Doubts, interpretation criteria, different points of view, etc. come into play and... other matters we all know about.

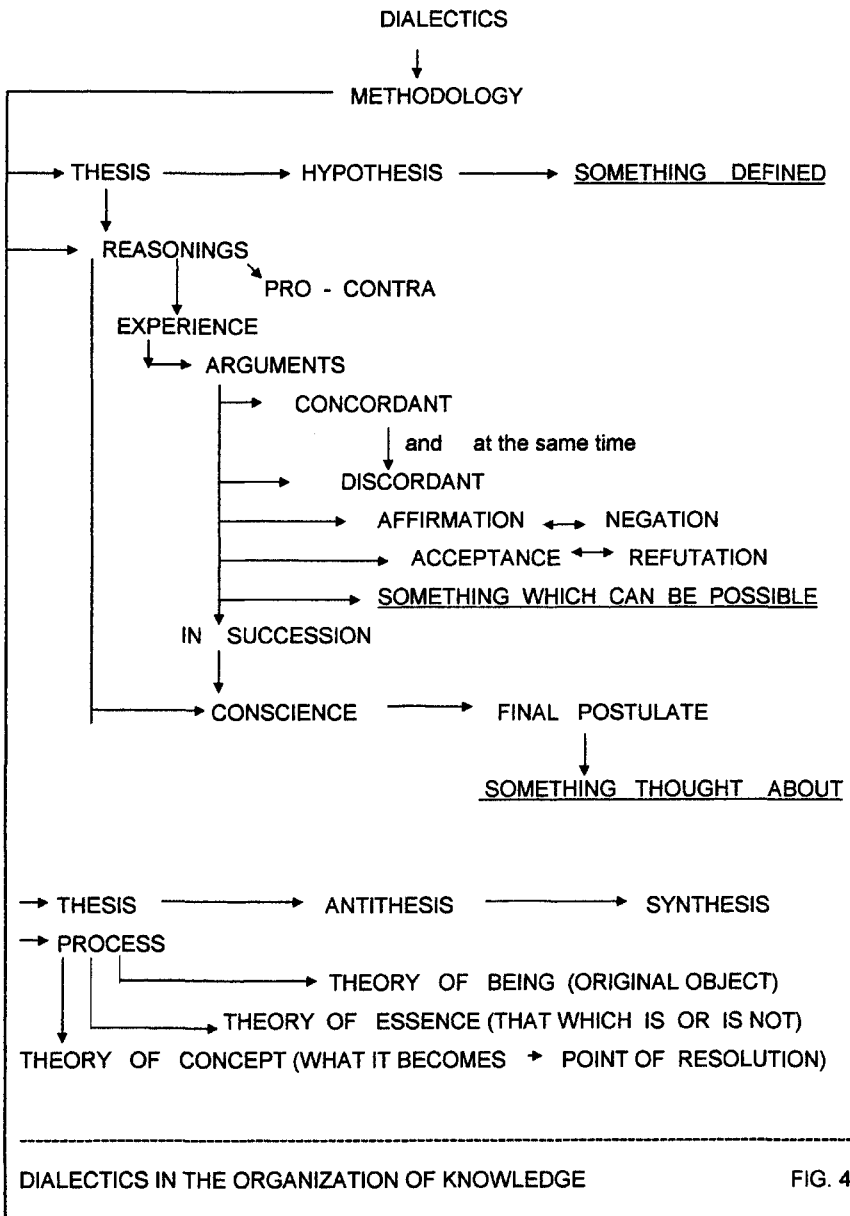
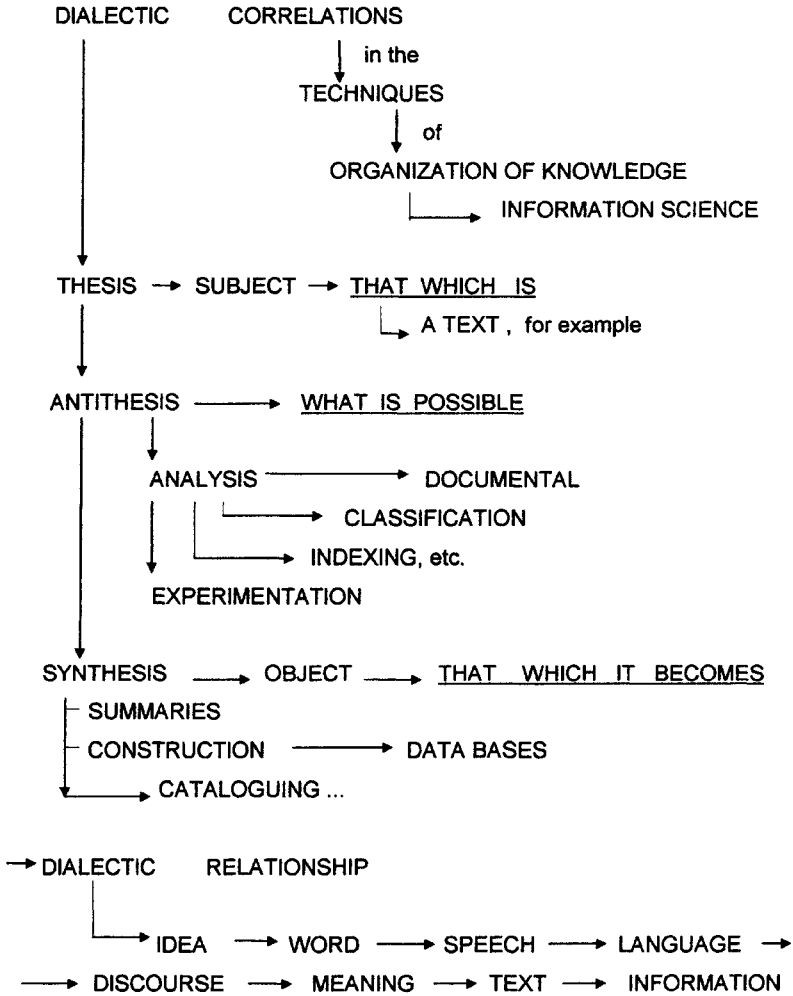


FIG. 4

DIALECTICS IN INFORMATION SCIENCE



Parmenides (n.f.s. 6 C B.C., Ferrater, 1981), to quote a classic scholar, also understands these matters and knows that after the ideas have been discriminated, one has to succeed in combining them again, without losing the meaning of the original ideas. Similarly, we can obtain other extrapolations from Hegel (1770-1831, Ferrater, 1981) who confirms the dialectic nature of information science techniques in the organisation of knowledge, when he defines concepts as: *subjective concept*, *objective concept* and *idea*. (Fig. 5)

A text provokes a new idea, with which the process starts all over again, not in spiral form, but in levels getting progressively higher. The new idea understands the previous one, and embraces it, giving it a greater degree of content. The successive leaps in time gradually mould knowledge and science into a single unit (Currás, 1993).

It is worth adding here the dialectic relationship between the idea, the word, speech, language, discourse, meaning, text, and information since all form an essential part of the actual content of the organisation of knowledge in information science. There is a continuous line running from the conception of the idea in the mind, originated by the information received, until we reach the text, and a new information. This line is never straight but shows a succession of ups and downs, following the same direction, which can be interrupted at any place and which ends there or carries on further. Each of these irregularities relates to dialectic moments, of successive and reciprocal concordances and discordances.

CONCLUSION

It would appear, from the writing in this paper, that we cannot refute the dialectic nature of knowledge organisation. A more detailed study will allow us to find a greater number of correlations and applications. In fact, the basic, fundamental principles have not changed from the times of the Greeks. What has changed is its concept as something extremely practical which can be applied, for example, to teaching and education, economic and political practices, socio-economic principles or something formal, forming part of logic or linking it to grammar and rhetoric, taking part in the forms of expression.

As we have been reasoning and endeavouring to show, dialectics is part of the empirical and practical development of knowledge organisation in its information science version and, therefore, we can draw up an approach to *information science dialectics*, or at least the nearest thing to it, which requires further in-depth study.

This may or may not be true. In these times of upheaval and transmutation, of uncontrollable discoveries and inventions, of profound changes in society, dialectics is applied to all branches of knowledge..... including everyday tasks. And if that were not enough, here it is applied to the organisation of knowledge, in the context of information science, in an approximation to *information science dialectics*.

REFERENCES

- Capone Braga, G., 1957: "Dialettica", in *Enciclopedia filosofica*, Venecia-Roma, Istituto per la Collaborazione Culturale (ed.), 1540-1559.
- Cole, Ch., 1994: "Operationalizing the notion of information as a subjective construct", *Journal of the American Society for Information Science*, 7 (45), 465-476.
- Currás, E., 1996: "Dialectic information systems and their connection with the new economy: some approaches to", in *Globalization of Information: Proceedings of 48th FID Congress*, Graz, FID (ed.), 61-68.
- Currás, E., 1996: *Tratado sobre ciencia de la información*, Rosario, Universidad Nacional de Rosario Editora.
- Currás, E., 1994: "Concepto de información en el contexto de integración científica", *AABADOM*, V (1), 4-8. "Pojem Informace v Kontextu Vedecké Integrace" (Concepto de información en el contexto de integración científica), *Kniznice a Informácie*, (Slovak Republic), 27 (10) 1995, 373-377; *Acta Bibliotecaria et Informatica*, (Czech Republic), 1996, 98-104.
- Currás, E., 1993: "The influence of systems on the concept of information", *ICS, Innovazione Comunicazioni Sviluppo*, V (3), 5-6; *Internat. Forum on Inf. and Doc.*, 18 (2), 32; "La influencia de la ciencia sistémica en el concepto de información", *Boletín de la Asociación de Bibliotecarios Profesionales de Rosario*, Segunda Epoca, 3; *Quím. e Ind.*, XLI (4) (1994), 49-51.
- Currás, E., 1993: "Vertical integration of science. A systemic approach" in *International systems science handbook: an introduction to systems for everybody*, Rodríguez Delgado, R., Banathy, B.H. (eds.), coedited by Currás, E., Madrid, Systemic Publications, chapter 3.7, 159-175.
- Currás, E., 1993: "Concierto y desconcierto en la organización del conocimiento actual y su intersección con el mundo de la información", *Ciencias de la Información*, 24 (6) (diciembre 1993), 210-225; *Scire*, 1 (1) (enero-junio 1995), 3-28.
- Currás, E., 1992: "La recuperación de la información como sistema abierto", in *Actas de las IV Jornadas de Documentación*, Barcelona, 363-375.
- Currás, E., 1992: "Information science-information as a dialectic interactive system" in *Cognitive Paradigms in Knowledge Organization*, Sarada Ranganathan Endowment for Library Science (ed.), 418-431; *IFID*, 20 (1) (January 1995), 31-42; "Información-ciencia de la información como sistema en interacción dialéctica", *Cuadernos de ADAB*, 1 (2 julio-dic. 1993), 366-380; *Referencias*, 1 (2) (marzo 1995), 9-17.
- Currás, E., 1990: "Dialectic interaction in science", *INICAE*, 9 (1) (March), 5-17.
- Currás, E., 1988: *La información en sus nuevos aspectos. Ciencias de la documentación*, Madrid, Paraninfo.
- *Enciclopedia filosófica*, 1996: Madrid, Grupo Zeta.
- Engels, F. (1820-1895), 1968: *Anti-dühring*, México, Ed. Grijalbo.
- Ferrater, J., 1981: *Diccionario de filosofía*, tomo I, A-K, Buenos Aires, Editorial Sudamericana.

- García Gutiérrez, A. G., 1996: *Procedimientos de análisis documental automático. Estudio de caso*, Sevilla, Consejería de Cultura, Junta de Andalucía, Instituto Andaluz del Patrimonio Histórico.
- García Gutiérrez, A. G., 1992: *Análisis documental del discurso periodístico* Madrid, Centro de Tratamiento de la Documentación.
- García Gutiérrez, A. G., 1990: *Estructura lingüística de la documentación: teoría y método*, Murcia, Universidad de Murcia.
- Meix, F., 1982: *La dialéctica del significado lingüístico*, Salamanca, Ediciones Universidad de Salamanca.
- Núñez Ruiz, D., 1975: "Dialéctica" in *Diccionario de Ciencias Sociales*, Madrid, Instituto de Estudios Políticos (ed.), 703-706.
- Pinto, M. ; Gálvez, C., 1996: *Análisis documental de contenido*, Madrid, Síntesis.
- Pinto, M., 1992: *El resumen documental. Principios y métodos*, Madrid, Fundación Germán Sánchez Ruipérez y Pirámide.
- Pinto, M., 1991: *Análisis documental. Fundamentos y procedimientos*, Madrid, Eudema.
- Rodríguez Delgado, R. (1912-1997), 1997: *Del universo al ser humano. Hacia una concepción planetaria para el siglo XXI*, Madrid, McGraw-Hill/ Interamericana de España.
- Sánchez Ortiz, R. ; Sánchez Ortiz, M., 1977: *Dialéctica*, Madrid, Ed. Bruño.
- Sandor, P., 1964: *Historia de la dialéctica*, Buenos Aires, Ediciones Siglo Veinte.
- Sichirollo, L., 1976: *Dialéctica*, Barcelona, Ed. Labor.
- Zeleny, J., 1982: *Dialéctica y conocimiento* Madrid, Ediciones Cátedra.

Madrid, September 1998

„KHSFlow“

Ein System zur hypertextbasierten Workflowmodellierung und –steuerung

Mohsen Shafaei

DATEV eG
Sigmundstraße 172, 90329 Nürnberg
email: Mohsen.Shafaei@datev.de

Inhalt

1. Einleitung
2. Das Konstanzer Hypertext System (KHS)
3. Das hypertextbasierte Workflow-System KHSFlow
 - 3.1 Einige Grundlagen
 - 3.2 Warum KHSFlow?
 - 3.3 Das KHSFlow-Workflowmodell
 - 3.4 Das Workflow-Beispiel „Bildungsurlaub“
 - 3.5 Der KHSFlow-Designer
 - 3.6 Die Architektur
4. Ausblick

Zusammenfassung

Workflow-Management-Systeme werden als Mittel zur Modellierung und Automatisierung betrieblicher Vorgänge (Workflows) eingesetzt. Dieser Beitrag zeigt, wie die Workflowmodellierung und -ausführung durch den Hypertext-Ansatz unterstützt werden kann. Im folgenden wird beispielhaft das hypertextbasierte Workflow-Management-System KHSFlow vorgestellt, welches auf das offene Hypertextsystem KHS (Konstanzer Hypertext System) aufsetzt. Es wird gezeigt, daß die Eigenschaften offener Hypertextsysteme dem Gegenstand des Workflow-Managements zugute kommen und sich damit einige der in diesem Bereich existierenden Probleme beseitigen lassen.

1. Einleitung

Workflows sind formal beschriebene Geschäftsprozesse, die von einem Workflow-Management-System zum Ablauf gebracht werden können. Ein Workflow-Management-System unterhält sowohl eine Modellierungskomponente, die der Spezifikation der Workflows dient, als auch eine Steuerungskomponente, die eine Ausführungsumgebung für Workflows definiert. Doch die meisten heute auf dem Markt vorhandenen Workflow-Management-Systemen beruhen auf herstellerspezifischen Modellierungs- und Steuerungskomponenten und zeigen immer noch Schwachstellen, vor allem bezüglich der Erweiterbarkeit und Wiederverwendbarkeit des Workflowmodells sowie Integration externer Informationsressourcen und -quellen. Es existiert keine einfache, einheitliche Workflowdefinitionssprache, die von jedem problemlos verstanden und eingesetzt werden kann. In der Workflowausführungsphase wird der Anwender oft zwischen verschiedenen Umgebungen bzw. Anwendungen geführt, die ihn bei der Workflowausführung unterstützen sollten. Dabei besteht aber die Gefahr, daß der Anwender schnell den Überblick verliert und sich überfordert fühlt.

Immer mehr Organisationen und Firmen erstreben eine unternehmensweit einheitliche Bürokommunikationsstrategie und versprechen sich durch den unternehmensweiten Einsatz von neuen attraktiven Technologien Zeit- und Kosteneinsparungen. Das Intranet ist ein Beispiel für solche Technologien. Als ein hypertextbasiertes unternehmensweites Netz integriert das Intranet die häufig heterogenen DV-Landschaften der Unternehmen unter der einheitlichen und benutzerfreundlichen Oberfläche eines WWW-Browsers. Der Einsatz von Workflow-Management-Systemen in verteilten heterogenen Umgebungen kann auch durch die Verwendung des WWW in Verbindung mit Komponenten, wie ActiveX und JavaBeans etc. unterstützt werden. Deswegen bemühen sich heute noch viele Hersteller der Workflow-Systemen um die Integration ihres Workflow-Systems mit dem WWW. Der Workflow-Pionier Action Technologies bietet beispielsweise mit Metro eine Webadaption an. Auch wenn nach Maßstäben der Hypertextforschung WWW es gar nicht verdient, als Hypertext-System bezeichnet zu werden, geht der Trend dahin Hypertext-Systeme mit dem WWW zu integrieren, um erstens die Inhalte der Hypertext-Systeme im WWW präsentieren zu können und zweitens intellektuelle, datenbankbasierte Hypertext-Systeme als WWW-Autoren-Systeme zu verwenden. Die Verbindung des offenen Hypertext-Systems KHS (Konstanzer Hypertext Systems) mit dem WWW ist ein Beispiel hierfür (vgl. Aßfalg 96). Hypertext-Systeme sind Informationssysteme, die seit Anfang der neunziger Jahre eine neue Forschungsrichtung innerhalb der Informatik etabliert haben. Sie finden ihre Anwendung vorwiegend in den Bereichen: Online-Dokumentation, Wörterbücher und Nachschlagewerke, Messen und Werbung, Journalismus, Fremdsprachen, Forschung, Wissensrepräsentation, Software Engineering, Workflow-Management (Prozeßanalyse), etc. Auch diese Informationssysteme können durch den Einsatz eines Workflow-Systems unterstützt werden. Nicht nur der Autor des Hypertextes könnte während seine Modellierungsarbeit (Aufbau des Hypertextes) unterstützt werden, sondern auch den Benutzern des Hypertextes kann dieser Einsatz zugute kommen. Daß die Verknüpfung der zwei Forschungsrichtungen Workflow-Management und Hypertext vorteilhaft ist, ist

offensichtlich. Wie sollen aber diese miteinander verknüpft werden? Kauft man sich ein Workflow-System, welches dann in eigenes Hypertextsystem integriert werden soll, oder integriert man besser ein Hypertextsystem in eigenes Workflow-System. Eine andere Vorgehensweise präsentiert der vorliegende Beitrag. Es wird für ein offenes datenbankbasiertes Hypertextsystem (das KHS), welches mit WWW verbunden ist und daneben auch auf andere Mächtigkeiten aufweist, eine Workflowkomponente konzipiert und implementiert. Es wird gezeigt, daß die Eigenschaften eines offenen Hypertextsystems bzw. die in so einem System bereits implementierten Funktionalitäten auch dem Workflow-Ansatz zugute kommen können und siehe einige der in diesem Bereich existierenden Probleme beseitigen lassen.

In den folgenden Abschnitten werden zuerst die grundlegenden Eigenschaften des zugrundeliegenden offenen Hypertextsystems KHS (Konstanzer Hypertext System) beschrieben. Danach wird das KHSFlow (hypertextbasierte Workflowkomponente des KHS) vorgestellt.

2. Das Konstanzer Hypertext System (KHS)

Ein Kriterium, das zur Kategorisierung der Hypertextsysteme führt, ist die Offenheit bzw. Geschlossenheit dieser Systeme gegenüber externen Informationsquellen und -ressourcen. Offene Hypertextsysteme ermöglichen die Einbindung externer Informationsquellen in ihre Hypertextbasis und bieten Funktionalitäten zur Integration externer Informationsressourcen an, während geschlossene Hypertextsysteme diese Funktionalitäten nicht vorsehen. KHS ist ein *offenes datenbankbasiertes Hypertextsystem*, an dem mehrere Benutzer gleichzeitig arbeiten können. KHS wurde am Lehrstuhl für Informationswissenschaft der Universität Konstanz unter der Leitung von Herrn Prof. Dr. Rainer Kühlen im Rahmen des DFG-Projekts WITH entwickelt. Die Implementierung des KHS erfolgte auf Unix-Workstations unter Einsatz der Smalltalk/VisualWorks Entwicklungsumgebung und des objektorientierten Datenbanksystems GemStone. Dank der Plattformunabhängigkeit der Smalltalk/VisualWorks ist KHS aber auch unter Macintosh, Windows95 und WindowsNT 4.0 lauffähig. KHS basiert auf die zuerst im [Kühlen95, S. 432] publizierten Eigenschaften für offene Hypertextsysteme und beruht auf der Typisierung von allen Hypertextobjekten. Hypertextobjekte sind Knoten und Links sowie Hypertexte selbst als Sammlung von Knoten und Links [Ham93b]. Knoten stellen inhaltstragende Elemente des Hypertextes dar, welche durch deren Typisierung verschiedenartige Informationen wie Fließtext, Abbildung und Text, Formulare, Tabellen, Audio Informationen, Graphen und zahlreiche anders formatierte Inhalte haben können. Jeder Knotentyp kann nur Informationen von bestimmter, vordefinierter Form und Struktur aufnehmen. Knoten können aber auch Anweisungen enthalten, die sofort ausgeführt werden, wenn der Knoten selektiert wird. Strukturknoten (non-terminale Einheiten) präsentieren im KHS weitere Knotentypen, die weitere terminale bzw. non-terminale Einheiten enthalten können und außerdem Gegenstand von Verknüpfungen sein können. KHS stellt eine Reihe von Interaktions- und Präsentationswerkzeugen zur Verfügung. Als wichtigstes Werkzeug ist der

KHS-Standardbrowser zu erwähnen, der dem Benutzer ständig zur Verfügung steht und ihn bei der Definition von Knoten und Verknüpfungen, bei der Präsentation und Modifikation der Knoteninhalte, bei der Navigation in den Verknüpfungsnetzen oder Strukturhierarchien, bei der Aktivierung weiterer Werkzeuge zum Retrieval sowie bei der Aktivierung umfassender (graphische) Übersichten etc., unterstützt. Das mit KHS abgedeckte Aufgabenspektrum geht weit über Standardaufgaben, wie man sie bei konventionellen Büroarbeitsplätzen findet, hinaus. Terminplanung, Adreßverwaltung, Literaturverwaltung, Korrespondenz (Email) mit Archivierung, Online-Recherche und Verwaltung der Ergebnisse, Erstellung von WWW-Hypertexten, Softwareerstellung, Diskussionforen und elektronische Zeitschriften sind während einer KHS-Sitzung keine Fremdwörter. Die Integration des KHS in den Internet-Mehrwertdienst WWW -über die direkte Anbindung der KHS-Datenbank und die dynamische Generierung der HTML-Dokumente zur Anforderungszeit- ergänzt die Offenheit des Systems (nach außen) und verleiht ihm einen extra Glanz. (Ham96) stellt unter anderem eine umfassende wissenschaftliche Ausarbeitung des KHS zur Verfügung.

3. Das hypertextbasierte Workflow-System KHSFlow

3.1 Einige Grundlagen

Jedes Workflow-Management-System muß die Mächtigkeit zur Erstellung, Modifikation und Interpretation von Workflowmodellen bieten. Dazu ist neben der Definition und Realisierung einer Sprache zur Formulierung dieser Modelle auch erforderlich, daß die Wartbarkeit und das Umfeld der Workflow-Management-Systeme berücksichtigt wird. Damit können die Modelle auch ausgetauscht und extern präsentiert werden. Viele Probleme bei der Umsetzung eines Anwendungssystems mit einem Workflow-Management-System sind auf ein unzulängliches Workflowmodell zurückzuführen. Jablonski stellt im (VoB96) eine Reihe fundamentaler Anforderungen an ein Workflowmodell vor, welche neben dem modularen Aufbau eines Workflow-Systems die wichtigsten Merkmale solcher Systeme darstellen. Er unterscheidet anwendungsbezogene Anforderungen (*Erweiterbarkeit, dynamische Anpaßbarkeit, Wiederverwendbarkeit und Offenheit*) von den modellbezogenen Anforderungen (*Graphische, formale und ausführbare Repräsentationsform*). Er sieht die Modularität als wichtigste Eigenschaft eines Workflowmodells an und führt unter Berücksichtigung der Unterscheidung zwischen sachlichen (funktionaler, operationaler, verhaltensbezogener, Informationsbezogener, organisatorischer und kausaler Aspekt) und technischen Aspekten (historischer und transaktionaler Aspekt) ein allgemeines aspektorientiertes Workflowmodell ein.

Auch die Architektur eines Workflow-Management-Systems muß einen modularen Aufbau aufweisen. Die WfMC (Workflow Management Coalition) stellt ein Standard-Referenzmodell vor, welches aus einem kontrollteil und fünf Schnittstellen (Prozeßdefinitionskomponente, Prozeßadministrations- und Überwachungskomponente, Applikations-komponente, Benutzerschnittstelle und Schnittstelle für Kommunikation mit anderen Workflowsystemen) besteht

und die Trennung der Komponenten für die Analyse und formale Beschreibung von Prozeßmodellen und dem Workflow-Kontrollteil (Engine) erzwingt.

3.2 Warum KHSFlow?

Betrachtet man das WfMC-Referenzmodell, die anwendungsbezogenen und modellbezogenen Anforderungen an ein Workflowmodell sowie die im (Kuhlen95) vorgestellten Eigenschaften der offenen Hypertextsysteme (OHS), so stellt man fest, daß die offene Hypertextsysteme, und somit auch das KHS, Eigenschaften aufweisen, die den Anforderungen an Workflowmodelle und Workflow-Management-Systeme gerecht werden. Diese Gemeinsamkeiten haben dazu beigetragen, daß man sich dafür entschieden hat, für das KHS eine eigene Workflowkomponente zu implementieren und kein auf dem Markt existierendes Workflow-System einzusetzen. Die folgende Tabelle faßt die wichtigsten Gemeinsamkeiten zusammen:

	Anforderungen an Workflowmodelle und -systeme	Eigenschaften der OHS
1	Erweiterbarkeit des Workflowmodells (neue Modellelemente)	OHS sind offen bezüglich des Umfangs (neue Knotentypen)
2	Wiederverwendbarkeit der Workflowspezifikationen	Typisierung der Knoten und Verknüpfungen
3	Offen bezüglich der Integration vorhandener und neuer Anwendungssysteme	Offen für heterogene Informationsressourcen
4	Bereitstellung der externen Daten	Einbindung externer Informationsressourcen
5	Schutz der Workflowtypen / Workflowinstanzen vor unbefugtem Zugriff	Datenbankorientierte / Filesystemorientierte Schutzmechanismen
6	Graphische Präsentation der Workflows	Diverse Präsentationwerkzeuge
7	Durchführung anderer Tätigkeiten neben dem Workflowgeschäft (andere Werkzeuge)	Komfortable Oberfläche, die dem Nutzer erlaubt, vom Hypertextsystem aus z.B. auch andere Mehrwertdienste des internationalen Informationsmarktes zu nutzen (E-Mail, Internet, Online-Datenbanken, Gopher, ...)

Das Projekt KHSFlow hat die prototypische Realisierung und Evaluierung eines hypertextbasierten Workflow-Systems am Beispiel personalbezogener Daten der DATEV eG zum Ziel und wird am Lehrstuhl für Informationswissenschaft der Universität Konstanz unter der Leitung von Herrn Prof. Dr. Rainer Kuhlen entwickelt. Die Implementierungsarbeiten erfolgen auf WindowsNT unter dem Einsatz der Entwicklungsumgebung VisualWorks3.0 (VisualWave30) und des relationalen Datenbanksystems Oracle. KHSFlow baut auf die bereits im KHS implementierten Funktionalitäten (Klassen und Methoden) und stellt zusätzliche Klassen und Funktionen zur Verfügung, die den Anwender in allen Phasen des Workflow-Managements (Informationserhebung, Analyse, Optimierung, Modellierung und Ausführung) unterstützen sollen.

3.3 Das KHSFlow-Workflowmodell

Ein Workflow kann anhand von Aspekten modelliert werden, die zusammen die verschiedenen Facetten eines Workflows beschreiben (BuJa94). Jablonski und Bußler stellen in (JaBu96) ein solches aspektorientiertes, modularaufgebautes Workflowmodell (MOBILE) vor, das im Rahmen der Workflowforschung am Lehrstuhl für Datenbanken der Universität Erlangen-Nürnberg entstand. Das Workflowmodell des KHSFlow ist ähnlich wie das MOBILE-Workflowmodell aufgebaut und berücksichtigt zur Zeit folgende Aspekte:

Der *funktionale Aspekt* definiert den Schritt bzw. die Schritte eines Workflows („Was“). Der *verhaltensbezogene Aspekt* beschreibt das „Wann“ der Ausführung. Hier wird die Abarbeitungsreihenfolge der Workflows unter Verwendung diverser Kontrollflußkonstrukten festgelegt. Der *organisatorische Aspekt* gibt Auskunft über die ausführenden Personen („Wer“, das kann auch ein Programm sein). Der *informationsbezogene Aspekt* spezifiziert die Daten, die bei der Ausführung der Workflows benötigt werden („Womit“, Datenfluß). Der *operationale Aspekt* beschreibt Ressourcen und Module, welche bei der Workflowsausführung benötigt bzw. gestartet werden („Wie“). Um die natürlichsprachliche Beschreibung der Workflows zu ermöglichen, sind bestimmte Modellelemente vorgesehen. Außerdem können weitere Beschreibungselemente jederzeit in das Modell integriert werden. Die Modellierung nach Aspekten erlaubt die Betrachtung von Workflows von einem bestimmten Gesichtspunkt aus und ermöglicht die Zerteilung des Modellierungsraumes in kleine, überschaubare Teile.

Im KHSFlow sind Workflows Instanzen der Klasse Workflowtyp. Eine Workflowinstanz ist entweder ein Elementar-Workflow oder ein Komposit-Workflow, welcher aus mindestens einem oder unter Umständen n Elementar-Workflows (Subworkflows) bestehen kann. Außerdem beinhaltet eine Workflowinstanz weitere Elemente, welche die oben erwähnten aspektbezogenen Informationen enthalten. Workflows werden im KHSFlow durch typisierte Strukturknoten dargestellt, die wiederum weitere Strukturknoten oder (terminale) Knoten enthalten können. Die Struktur dieser Knoten ist ausgehend von dem aktuellen Workflowmodell vorgegeben und kann jederzeit erweitert bzw. angepaßt werden. Um das Workflowmodell zu veranschaulichen, wird im folgenden zuerst ein Beispiel-Workflow modelliert.

3.4 Das Workflow-Beispiel „Bildungsurlaub“

Ein Bildungsurlaub (Kurz: Bur) muß vom Mitarbeiter schriftlich beantragt werden (BurBeantragen). Nach der Genehmigung des Antrages (BurGenehmigen) durch den jeweiligen Vorgesetzten wird der Antrag entweder abgelehnt (BurAblehnen) oder an die Personalabteilung zur Prüfung (BurPrüfen) weitergeleitet. Dort wird zuerst geprüft, ob das Unternehmensgesamtkontingent für Bildungsurlaube bereits ausgeschöpft ist (GesKontPrüfen) oder ob der Antragsteller sein persönliches Bildungsurlaubskontingent (maximal sechs Tage innerhalb von zwei Jahren) erreicht hat (PersKontPrüfen). Abhängig von dieser Prüfung bekommt der Antragsteller automatisch entweder eine Absage

(BurAbsage) oder eine Zusage (BurZusage). Im Falle einer Zusage werden die Antragsdaten an das Gleitzeitsystem weitergeleitet (S. Abb. 1)

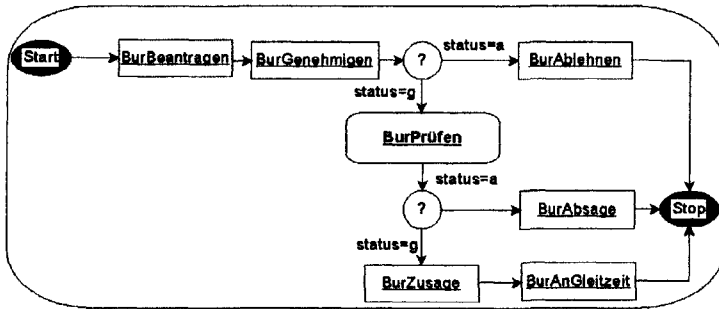


Abb. 1 : Workflowtyp „Bildungsurlaub“

Der funktionale Aspekt ist in der Abbildung durch Rechtecke (elementare Workflows) bzw. Rechtecke mit gerundeten Ecken (kompositere Workflows) beschrieben. Der Workflowtyp „Bildungsurlaub“ ist also ein kompositere Workflow, der aus sechs elementaren und einem kompositere Workflow (BurPrüfen) besteht. Der Subworkflow „BurPrüfen“ ist wiederum ein kompositere Workflow, der aus zwei elementaren Workflows „GesKontPrüfen“ und „PersKontPrüfen“ besteht (graphisch nicht dargestellt). Der verhaltensbezogener Aspekt ist durch Kontrollkonstrukte und Pfeile zwischen den einzelnen Workflows dargestellt. Die von Workflows benötigten Daten, die ausführende Personen der Workflows sowie die Programme (Funktionen bzw. Module), die mit den elementaren Workflows verbunden sind, sind hier nicht dargestellt, müssen jedoch während der Modellierungsarbeit spezifiziert werden. Um das Beispiel-Workflow mit KHSFlow zu modellieren, geht man wie folgt vor:

1. Man erstellt bzw. öffnet ein Hypertextdokument, welches mehrere typisierte Hypertextobjekte enthalten kann (auch Workflowtypen sind Hypertextobjekte) und
2. fügt in das Hypertextdokument einen Knoten vom Typ Composit-Workflow ein. Das System fragt nach dem Knotennamen (in unserm Fall Bildungsurlaub) und fügt den Knoten ein, welcher automatisch weitere (vorher definierte) aspektspezifische Sub-Knoten (Description, Variablen, Constraints, ControlFlow, DataFlow und Rolle) beinhaltet. In den Knoten **Description** kann man die verbale Beschreibung des Workflows eingeben. Der komposite Knoten **Variablen** kann weitere Knotentypen (Input,Output,Inout) enthalten und spezifiziert Workflow-Variablen (z.B. Input initiator: String oder Input antrag: Antrag). Hier können die bekannten Datentypen (String, Integer, Float,...) sowie vorher definierte Datentypen bzw. Datenstrukturen verwendet werden. Der komposite Knoten **Constraints** kann ebenso weitere spezielle Knoten (EnterConstraint, RunConstraint sowie ExitConstraint) enthalten, die beispielsweise angeben, wann ein Workflow nicht ausgeführt werden darf (Beispiel: RunConstraint (not after 20pm). In den Knotentyp **Rolle** kann man beispielsweise die wichtigsten Daten (Name, Vorname, Abteilung, Email-Adresse Zuständigkeit,...) einer Person abspeichern, die

den Workflow ausführt. Im Knoten **ControlFlow** wird festgelegt, in welcher Reihenfolge (bei welchen Zuständen) die Workflows ausgeführt werden müssen. Schließlich kann der Datenfluß zwischen den Workflows in den kompositen Knoten **DataFlow** spezifiziert werden.

3. Als nächster Schritt muß der Knoten Bildungsurlaub selektiert werden (man befindet sich in diesem Knoten), um weitere Workflow-Knoten aus unserem Beispiel einfügen zu können (beliebige Erfassungsreihenfolge). Ein elementarer Workflow-Knoten beinhaltet ebenso weitere Knotentypen (Description, Variablen, Constraints, Application und Rolle). Der Knotentyp **Applikation** beinhaltet die Operation (Funktion, Methode, Modul, Programm), die aktiviert wird, wenn ein elementarer Workflow ausgeführt wird (also das „Wie“). Somit ist die Struktur des kompositen Workflows „Bildungsurlaub“ gestaltet. Die Abbildung 2 zeigt den KHS-Standardbrowser nach der Selektion des Knotens Bildungsurlaub.

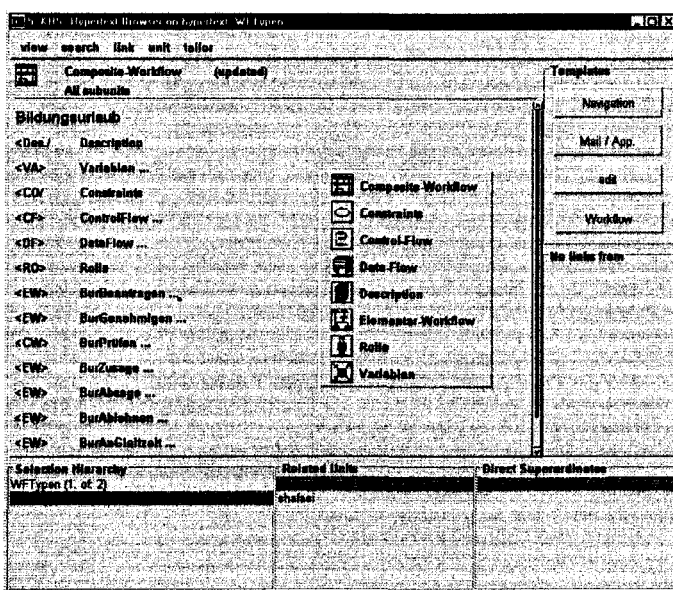


Abb. 2: Der KHS-Standardbrowser. Selektiert ist der komposite Workflow-Knoten „Bildungsurlaub“. Das mittlere Fenster zeigt alle Knotentypen, die in den aktuell selektierten Knoten einfügbar sind.

3.5 Der KHSFlow-Designer

KHSFlow stellt zusätzlich zum KHS-Standardbrowser einen speziellen Workflowmodellierungsbrowser (den KHSFlow-Designer) zur Verfügung, der dem Benutzer die Möglichkeit bietet, Workflows via Drag&Drag zu spezifizieren. Die folgenden Abbildungen zeigen den Designer bei der Darstellung des kompositen Workflow-Beispiels „Bildungsurlaub“ und des elementaren Workflows

40



Der Mausklick auf ein Listenelement verursacht die sofortige Navigation zu Strukturknoten, der Sub-Knoten enthält, so entsteht automatisch ein solches Netz, wie es im mittleren Fenster der Abbildungen (Abb. 3) zu sehen ist (der Doppelklick auf ein Rechteck (CompositeUnitFigur/FormUnitFigur) bewirkt das gleiche), sonst wird der Inhalt des elementaren Knotens dargestellt. Rechtecke präsentieren Knoten (Ikone betonen den Knotentyp) und Pfeile bringen die referentiellen Knoten-Verknüpfungen zur Darstellung. Der Knotentyp ControlFlow enthält die strukturellen Verknüpfungen, die die Abarbeitungsreihenfolge von Knoten (Workflows) bestimmen. Wie in der Abbildung 4 ersichtlich (s. Toolbar), werden hier bestimmte Kontrollflußkonstrukte (Sequence, Parallel, Bedingte Ausführung, Limitierung, Verzögerung, And, Or, Start, Stop und Auswahl) sowie spezielle Mechanismen zur Verfügung gestellt, die dem Modellierer erlauben, Kontrollflüsse in gewohnter Form, also via Drag&Drop und per Mausklick, spezifizieren zu können. Die Abbildung zeigt die Kontrollflüsse des Workflow-Beispiels „Bildungsurlaub“. Um die Parallelität zur Darstellung zu bringen, wurde das Beispiel leicht modifiziert (Parallele Ausführung der Workflows „BurZusage“ und „BurAnGleitzzeit“).

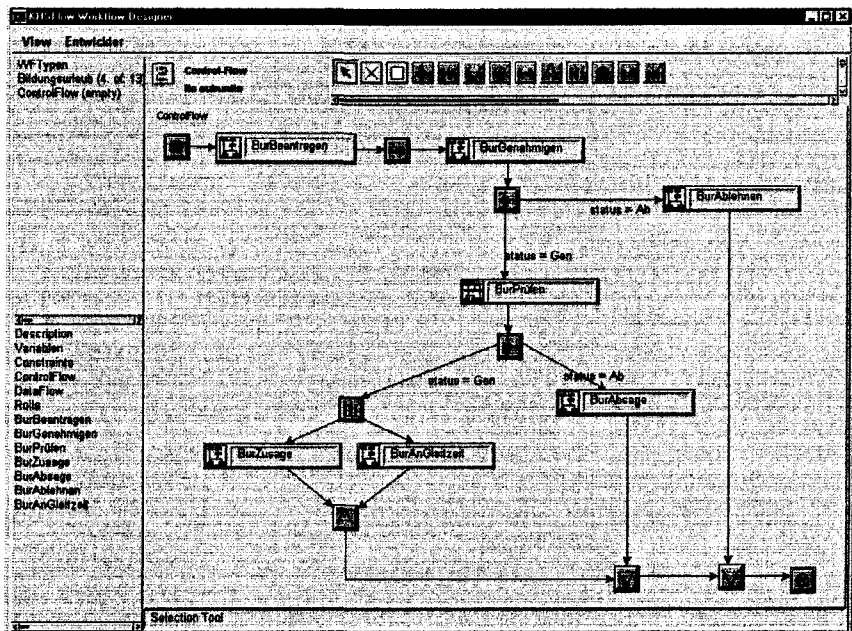


Abb. 4: Der KHSFlow-Designer. Selektiert ist der Knotentyp „ControlFlow“, der die aktuelle Abarbeitungsreihenfolge des kompositen Workflows „Bildungsurlaub“ graphisch darstellt.

Das grob vorgestellte KHSFlow-Workflowmodell kann mit geringem Aufwand auf weitere Elemente erweitert werden. Ein wesentlicher Vorteil des Modells

besteht darin, daß der Modellierer keine besonderen Kenntnisse bezüglich einer Scriptsprache haben müßte, um Workflowtypen definieren zu können. Die definierten Workflowtypen können samt ihrer Struktur und Subworkflows innerhalb eines anderen Workflowtyps wiederverwendet werden. Zu beachten ist noch die Tatsache, daß der hier eingeführte Strukturknoten „ControlFlow“ (ControlConstructions) nichts anderes darstellt als spezielle Hypertext-Verknüpfungen, die nicht nur genau zwei Hypertextobjekte miteinander verknüpfen, sondern auch die Verbindung von mehreren Ursprungs- und Zielobjekten miteinander erlauben.

Nachdem das KHSFlow-Workflowmodell sowie der KHSFlow-Designer grob eingeführt sind, soll in den folgenden Abschnitten die Architektur des KHSFlow kurz dargestellt werden. Danach werden die wichtigsten Komponenten des KHSFlow aufgelistet.

3.6 Die Architektur

Eine grobe Betrachtung des Systems läßt erkennen, daß das System aus zwei Teilen besteht: Server-Teil und Client-Teil. Der Server ist eine leistungsfähige Maschine, auf welcher die KHS-Anwendung und somit auch das KHSFlow ständig aktiv sind. VisualWorks 3.0 und VisualWave3.0 bilden die Basis für das KHS und KHSFlow. Der Server ist fähig, sowohl mit der objektorientierten Datenbanksystem GemStone als auch mit der relationalen Datenbanksystem Oracle zu kommunizieren. Auch Zugriffe auf Hostdatenbanken (wie z.B. DB2-Datenbanken) sind vom Server aus (über ODBC-Schnittstellen) möglich. Bedingt durch den Einsatz vom VisualWave kann der Anwender sowohl über das Intranet (unternehmensintern) als auch über das Internet (unternehmensübergreifend) auf die im KHS bzw. KHSFlow implementierten Anwendungen zugreifen. Der Anwender muß hierfür natürlich ein Zugriffsrecht haben (d.h. dem Server bekannt sein). Wir werden in den folgenden Abschnitten sehen, wie das funktionieren kann. Abbildung 5 stellt die Umgebung des KHSFlow ganz grob dar. KHSFlow ist vorerst als ein in diese Umgebung integriertes Workflow-System zu betrachten. KHSFlow ist modular aufgebaut, basiert auf das WfMC-Referenzmodell und besteht aus sieben Komponenten, die im folgenden grob erläutert werden.

KHSFlow-Designer. Die Workflowtypen (Modelle) können via KHS-Standardbrowser definiert werden und mit Hilfe des Table-Of-Content-Browser des KHS in grafischer Form betrachtet werden. Der KHSFlow-Designer verknüpft die Funktionalitäten der beiden Browser. Während der Table-Of-Content-Browser nur die graphische Präsentation der definierten Modelle ermöglicht, bietet der KHSFlow-Designer eine graphische Oberfläche zur Definition, Manipulation und Präsentation der Modelle.

KHSFlow-Admin. Diese Komponente unterstützt den Administrator des Systems bei Administration und Überwachung der Workflows und arbeitet eng mit dem Kontrollteil, der u. a. auch über die Stati der Workflows Bescheid weiß, zusammen.

KHSFlow-OrgManager. Die Historie und somit auch die aktuelle Organisationsstruktur des Unternehmens (Aufbauorganisatoin) ist hier abgebildet. Der KHSFlow-OrgManager liefert auf Anfragen des Kontrollteils bzw. des KHSFlow-Designers Informationen über den Bearbeiter des Workflows.

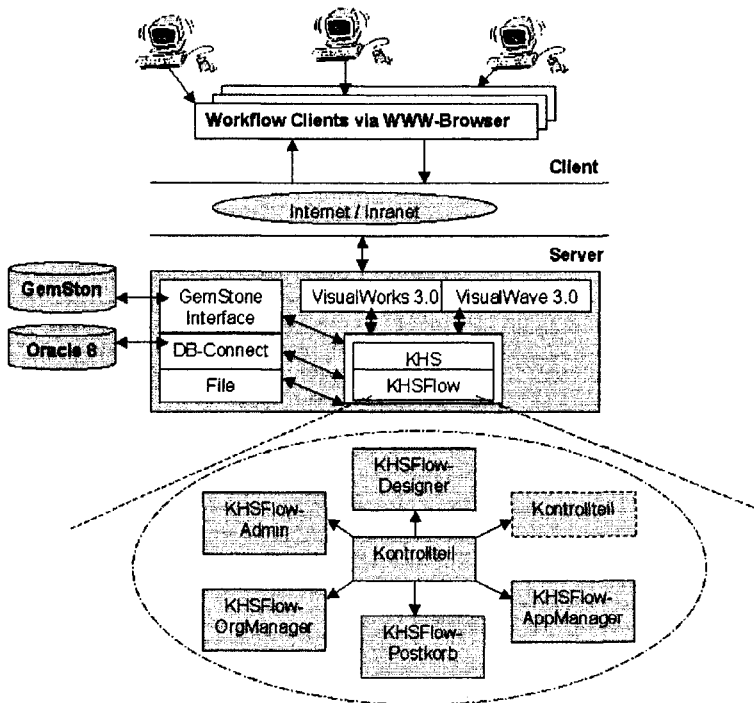


Abb. 5: Die Architektur und wichtigste Komponenten des KHSFlow

KHSFlow-AppManager. Spezielle Applikationsaufrufe, Java-Applets, ActiveX-Komponenten und Smalltalk-Controller, die Methoden zum Aufruf bestimmter Applikationen enthalten oder selbst Applikationen öffnen und verwalten, sind dem ApplikationManager bekannt und werden beim Bedarf vom ihm gestartet bzw. referenziert.

KHSFlow-Postkorb. Der KHSFlow-Postkorb ist die Schnittstelle des Endanwenders (Client) zum KHSFlow. Der Postkorb ist ein speziell für das KHSFlow entwickelter Dialog, der bedingt durch den Einsatz von VisualWave auch webfähig ist. Der Anwender kann den KHSFlow-Postkorb vom WWW-Browser aus aufrufen (Der Postkorb kann auch bei dem Client direkt installiert werden). Nach der Identifikation und Authentifikation des Benutzers wird dann seine Arbeitsliste aufgebaut. Der Anwender kann Workflows, die auf seiner Liste existieren, bearbeiten, weiterleiten, löschen, etc. Aus dem aktuellen Workflowtyp-Katalog,

der ebenso hier zur Verfügung steht, kann der Anwender Workflows instantiieren (s. Abb. 6).

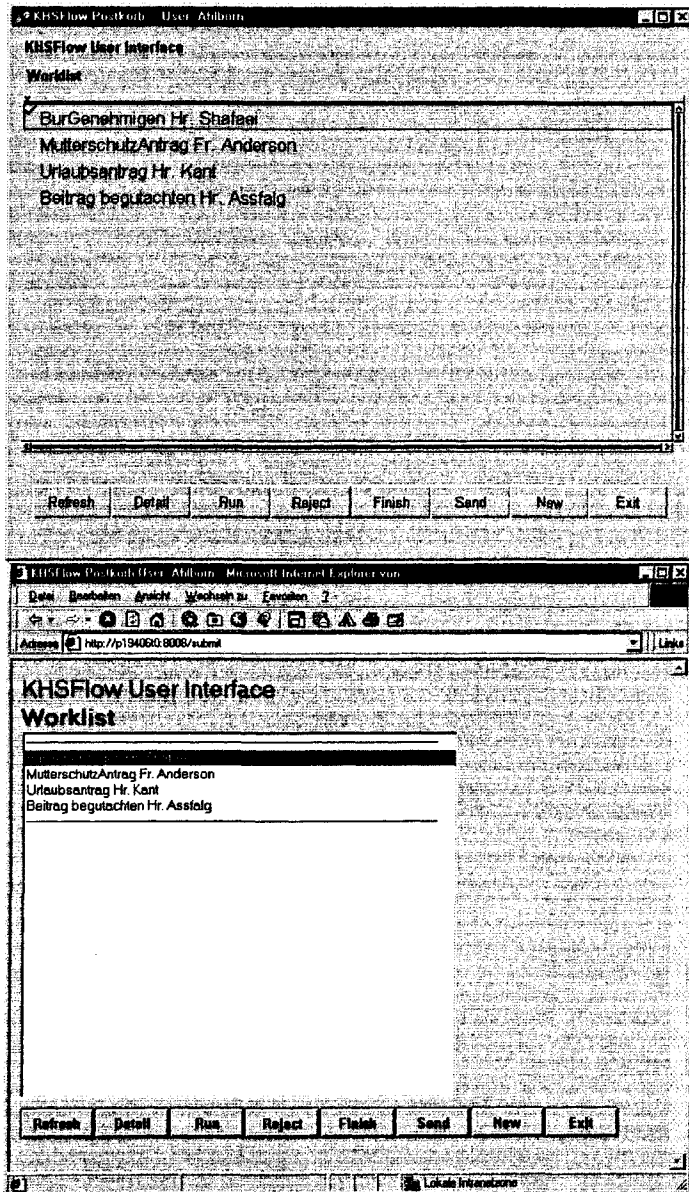


Abb. 6: Der KHSFlow-Postkorb

Der Kontrollteil. Der Kontrollteil ist das Herz des KHSFlow. Eine der wichtigsten Aufgaben des Kontrollteils ist die Instantiierung des abstrakt definierten Workflowtypen. Auf Grundlage des aktuellen Zustands eines Vorgangs sowie des vorgegebenen abstrakten Modells entscheidet der Kontrollteil, welcher Workflow als nächster zu bearbeiten wäre. Der Kontrollteil besteht aus mehreren Modulen (z.B. ObjectMaker, WorkTable, StatusManager,...). Der Kontrollteil kommuniziert ständig mit anderen Komponenten, die für ihn bestimmte Aufgaben erledigen.

4. Ausblick

Wir arbeiten momentan an der Implementierung des oben vorgestellten Ansatzes. Zur Zeit können Workflows mit Hilfe des KHSFlow-Designers problemlos modelliert werden. Die aktuell verfügbaren Kontrollflußkonstrukte ermöglichen einfache Modellierung komplexer Abhängigkeiten. Neue Kontrollflußkonstrukte können problemlos eingeführt werden. Obwohl einige Komponenten des Systems noch nicht (fertig) implementiert sind, können zur Zeit doch einfache Workflows ausgeführt werden. Der Client kann von WWW-Browser aus mit dem Server kommunizieren und Workflows bearbeiten oder instantiieren (starten). Die Zusammenarbeit zwischen KHSFlow-Komponenten läuft befriedigend. Parallel zur Weiterentwicklung des Systems, modellieren wir die Geschäftsprozesse (Workflows) der Personalabteilung der DATEV mit Hilfe des KHSFlow. Dadurch wird die Funktionalität des Systems anhand praxisrelevanter Beispiele getestet und bei Bedarf erweitert.

Literatur

- [Aßfal96] R. Aßfal: *Integration eines offenen Hypertext-Systems in den Internet-Mehrwertdienst World Wide Web*. Ein Ansatz unter Verwendung eines objektorientierten Datenbanksystems, Hartung-Gorre Verlag, Konstanz 1996.
- [BuJa94] Bußler C., Jablonski S.: An Approach to Integrate Workflow Modeling and Organization Modeling in an Enterprise. In: Proc. 3rd IEEE International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprise (1994), S. 81-95.
- [Conklin87] J. Conklin: *Hypertext: An introduction and survey*, IEEE Computer 20 (1987) 9. S. 215-220.
- [HaRi93] Hammwöhner R., Rittberger M. : *KHS - Ein offenes Hypertext-System*, Bericht 28-93 (WITH-3/93) Mai 1993, Informationswissenschaft, Universität Konstanz.
- [Ham93b] Hammwöhner R. , Kuhlen R. : *Semantik Control of Open Hypertext Systems by Typed Objects*, Bericht 34-93 (WITH-6/93) August 1993, Informationswissenschaft, Universität Konstanz.
- [Ham96] Hammwöhner R. : *KHS - Ein offenes Hypertext System*, Habilitation 1996, Informationswissenschaft, Universität Konstanz.

- [Jab95] Jablonski S. : *Workflow-Management-Systeme, Modellierung und Architektur*. Thomson's aktuelle Tutorien, TAT 9, International Thomson Publishing, Bonn 1995.
- [JaBu96] Jablonski S., Bußler C. : *Workflow Management Modeling Concepts, Architecture and Implementation* International Thomson Computer Press, London, Bonn,... 1996.
- [Kuhlen91] Kuhlen R.: *Hypertext - Ein nicht-lineares Medium zwischen Buch und Wissensbank*, Edition SEL-Stiftung, Springer-Verlag, Berlin 1991.
- [Kuhlen95] Kuhlen R.: *Informationsmark, Chancen und Risiken der Kommerzialisierung von Wissen*, Schriften zur Informationswissenschaft Band 15, Univ.- Verlag, Konstanz 1995.
- [Schulz97] Schulze W.: *Workflow-Management Entwicklung von Anwendungen und Systeme*, dpunkt.verlag, Heidelberg 1997
- [Shafaei96] Shafaei M.: *Workflow & Hypertext*, Diplomarbeit 1996, Informationswissenschaft, Universität Konstanz.
- [ÖstV96] Österle H., Vogler P.: *Praxis des Workflow-Managements, Grundlagen, Vorgehen, Beispiele*, Vieweg, Wiesbaden 1996.
- [VoB96] Vossen G., Becker J.: *Geschäftsprozeßmodellierung und Workflow-Management: Modelle, Methoden, Werkzeuge*, 1. Auflage, Thomson Publishing GmbH, Bonn 1996.

Entscheidungsunterstützung bei der Planung von Übersetzungsprojekten

Rainer Hammwöhner

Philosophische Fakultät IV - Sprach- und Literaturwissenschaft
Lehrstuhl für Informationswissenschaft
Universität Regensburg
93040 Regensburg
email: Rainer.Hammwoehner@sprachlit.uni-regensburg.de

Inhalt

- 1 Einführung
- 2 Entscheidungsparameter in Übersetzungsprojekten
- 3 Präzisierung der Projektziele und Skizzierung eines Entscheidungsmodells
- 4 Kosten / Nutzen
5. Ausblick

Zusammenfassung

In diesem Beitrag werden die Ziele des von der EU geförderten Projekts Trans-Router vorgestellt, in dessen Rahmen ein Entscheidungsunterstützungssystem für Übersetzungsagenturen und –dienste entworfen und prototypisch entwickelt werden soll. Das System soll den Manager bei der Auswahl der Übersetzungsverfahren und –hilfsmittel für konkrete Übersetzungsprojekte unterstützen. Dabei werden Eigenschaften der zu übersetzenden Texte, der zur Verfügung stehenden Ressourcen und der Übersetzungsprojekte in die Entscheidung einbezogen¹.

¹ Ich möchte mich an dieser Stelle bei Nancy Underwood und Bart Jongejan (CST), Keith Brazil und Conor McDonagh (LRC), Jo Calder (LTG), für die anregende Diskussion während des letzten Projektmeetings bedanken. Die in Abb. 1 skizzierten Profile sind auf einen bisher nicht publizierten Entwurf von Nancy Underwood zurückzuführen.

1 Einführung

Große Übersetzungsagenturen und -dienste werden mit einer großen Bandbreite von Übersetzungsprojekten konfrontiert. Die Projekte unterscheiden sich in der Länge und Komplexität der zu übersetzenden Texte, dem Fachgebiet, der erwarteten Übersetzungsqualität und den einzuhaltenden Terminen. Um jedes Projekt erfolgreich, aber auch effizient abschließen zu können, ist es erforderlich, den Übersetzungsvorgang so zu planen, daß die vorhandenen Ressourcen so optimal wie möglich eingesetzt werden (Swanholm 92). Als Ressourcen stehen der Agentur einerseits qualifizierte Übersetzer zur Verfügung und andererseits computergestützte Übersetzungssysteme und Übersetzungshilfen wie Wörterbücher, Translation-Memories etc., die teilweise zu kohärenten Arbeitsumgebungen zusammengefaßt sind (Le-Hong et. al 92). So lassen sich unter Einsatz eines Übersetzungssystems² umfangreiche Texte vergleichsweise schnell und kostengünstig übersetzen. In diesem Fall dürfen aber nicht zu hohe Anforderungen an die Qualität gestellt werden. Qualifizierte Übersetzer hingegen sind teuer und können nicht so große Textmengen bewältigen. Die Entscheidung über die in einem konkreten Projekt zu wählende Vorgehensweise fällt der Translation-Manager. Er steht vor der Aufgabe, einerseits die relevanten Eigenschaften eines Textes einzuschätzen und daraus einen Arbeitsablauf unter Einschluß der einzusetzenden und eventuell noch zu entwickelnden oder anzuschaffenden Ressourcen zu entwerfen. Daraus folgt, daß der Translation-Manager einen Überblick über die auf dem Markt befindliche und die in der Agentur unmittelbar verfügbare Software haben muß, sowie die einsetzbaren Übersetzer und ihre individuellen Stärken und Schwächen. Als besonders schwierig erweist sich auch eine angemessene Erfassung und Einschätzung von Übersetzungsressourcen, die im Zusammenhang mit abgeschlossenen Projekten entstanden sind. Bestehende Translation-Memories werden z.T. nicht wieder aufgefunden und sind nicht leicht hinsichtlich ihres Nutzens für ein neues Projekt einzuschätzen. Von Umfang und Qualität dieses Erfahrungswissens eines Translation-Managers hängt die Arbeitsplanung einer Übersetzungsagentur und damit die Qualität der Dienstleistung entscheidend ab.

Ziel des Projekts TransRouter ist es, dem Translation-Manager in diesem komplexen Auswahlprozeß durch ein Entscheidungsunterstützungssystem (für eine Einführung siehe (Turban 95)) Hilfestellung zu geben, das während des Projekts als Prototyp entwickelt werden soll. Bei der Entwicklung dieses Prototypen sind dabei folgende Fragen zu klären:

Was soll der Gegenstand der Beratung sein? Zu den Planungsproblemen innerhalb eines Übersetzungsprojekts gehört nicht nur die Auswahl von Übersetzungsverfahren sondern auch die Reservierung der jeweils benötigten Ressourcen im Sinne eines Produktionsplanungsprozesses.

² Für eine Einführung in maschinelle Übersetzung siehe (Hutchins; Somers 92).

Wie weitgehend soll die Beratungsleistung sein. Sollen Werkzeuge oder Vorgehensweisen bewertet werden oder soll das System eigenständig entscheiden?

Welche Eigenschaften von Texten, Projekten und Übersetzungsressourcen sind relevant für die Auswahl des Übersetzungsverfahrens?

Wie können diese Eigenschaften intellektuell oder automatisch erhoben werden?

Welche regelhaften Zusammenhänge erlauben die Auswahl von Übersetzungsverfahren anhand der Eigenschaften der Projekte, Texte und Ressourcen?

Die Beantwortung der oben aufgeführten Fragen, insbesondere nach relevanten Texteigenschaften und regelhaften Zusammenhängen, erfordert Expertise aus konkreten Übersetzungsprojekten. Deshalb ist im Projekt TransRouter viel Raum für die Konsultation von Übersetzungsfachleuten und –agenturen vorgesehen.

Die Projektpartner entstammen dem Anwenderkreis aus dem Bereich Übersetzung und Lokalisierung (Berlitz Ireland Ltd, Dublin; Localization Resources Center, LRC, Dublin) und sprach- und informationswissenschaftlichen universitären Einrichtungen (Center for Sprogteknologi, CST, Kopenhagen; Institut Dalle Molle pour Etudes Sémantiques et Cognitives, ISSCO, Genf; Language Technology Group, LTG, Edinburgh; Lehrstuhl für Informationswissenschaft der Universität Regensburg). Mit der Gesellschaft für multilinguale System, GMS, München ist ein Anbieter von Übersetzungssoftware vertreten.

2 Entscheidungsparameter in Übersetzungsprojekten

Die Zuteilung von Ressourcen in Übersetzungsprojekten ist abhängig von Parametern, die sich unmittelbar dem Projekt zuordnen lassen oder Eigenschaften der Übersetzungsagenten bzw. -ressourcen erfassen (s.a. Abb. 1).

2.1 Projektparameter

Ein Übersetzungsprojekt ist durch vielfältige Parameter gekennzeichnet, von denen einige aus den Eigenschaften der zu übersetzenden Text resultieren. Auf diese wird im nächsten Abschnitt eingegangen. Die Texteigenschaften sind allerdings nur im Zusammenhang mit weiteren durch das Projekt definierten Rahmenbedingungen zu interpretieren. Eine hohe Komplexität des zu übersetzenden Textes schließt eine maschinelle Übersetzung nur im Zusammenhang mit bestimmten Qualitätsvorgaben aus. Qualität ist wiederum ein komplexes Merkmal, daß Eigenschaften wie Fehlerhäufigkeit, Verständlichkeit und Stil umfaßt (O'Sullivan 92).

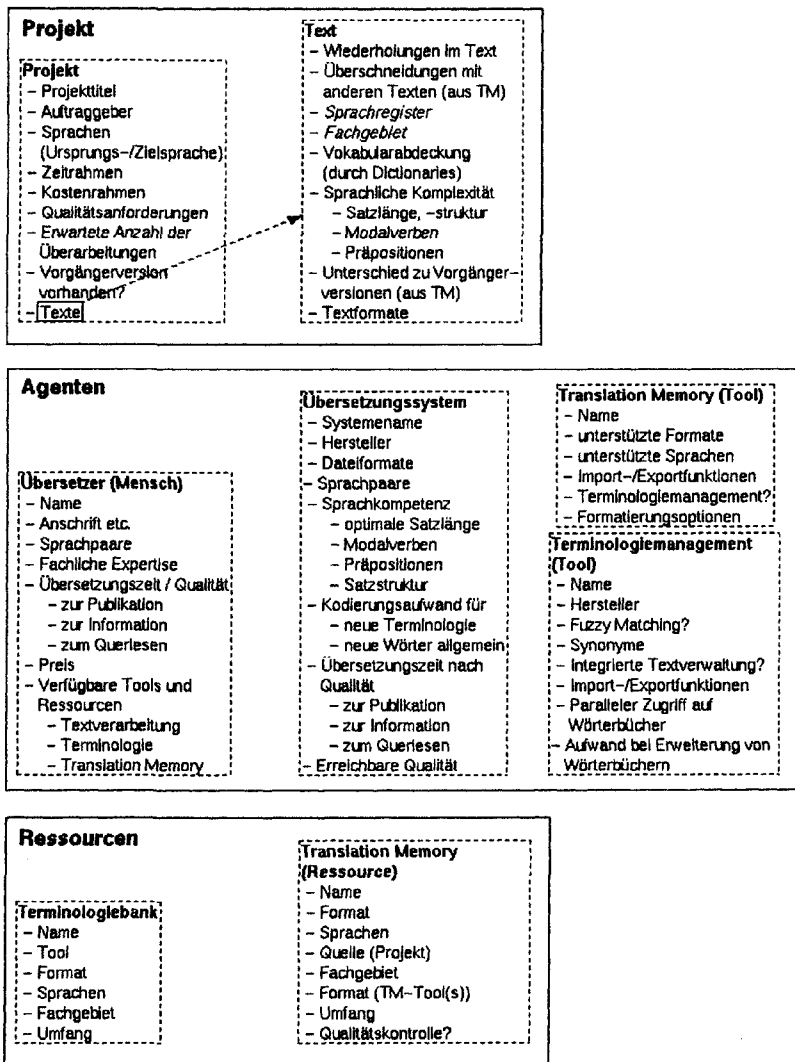


Abb. 1: Profile von Übersetzungsprojekten, -agenten und -ressourcen

Im Zusammenhang mit TransRouter wird zwischen drei Qualitätsstufen unterschieden:

Publikationsqualität setzt einen guten Stil in Kombination mit einer hohen Verständlichkeit und einer sehr geringen Fehlerrate voraus.

Informationsvermittlungsqualität stellt geringere Ansprüche vor allem an den Stil. Die tolerierbare Fehlerrate ist an der Verständlichkeit zu messen, bei der allenfalls geringe Abstriche erlaubt sind.

Browsingqualität erlaubt ein Querlesen des übersetzten Textes, so daß ein Eindruck von der Thematik und dem Ziel des Textes entsteht. Für eine tiefere Würdigung ist dann eine höherqualitative Übersetzung anzufertigen.

Weiterhin sind für ein Projekt Zeit- und Kostenrahmen von Bedeutung. Wichtig für Übersetzungsprojekte ist, daß viele große Organisationen weitgehend normierte Texte in regelmäßigem Turnus übersetzen lassen müssen (Vierteljahresberichte, Handbücher etc.). Bei derartigen Wiederholungen kann eine Investition in Ressourcen wie elektronische Wörterbücher oder Translation-Memories sinnvoll sein. Ist das aktuelle Projekt nicht das erste in einer solchen Sequenz, muß der Zugriff auf bestehende Ressourcen gewährleistet werden. Diese sind dann dahingehend zu beurteilen, ob die Abdeckungsquote für die neue Fassung des Textes zufriedenstellend ist, oder ob eine völlige Neuübersetzung vorzuziehen ist.

2.2 Texteigenschaften

Bei der Beurteilung von Texten hinsichtlich der anwendbaren Übersetzungsverfahren sind sprachliche Eigenschaften, wie Wortschatz oder syntaktische Komplexität, und solche der Textrepräsentation, wie Formatierung oder Speicherformat, zu berücksichtigen. Während die sprachlichen Kriterien unmittelbaren Einfluß auf die Komplexität des Übersetzungsprozesses und die damit in Frage kommenden Hilfsmittel haben, sind diejenigen der Textrepräsentation nur mittelbar relevant. Sie können bestimmen, ob ein konkreter Text von einem Übersetzungsprogramm verarbeitet werden kann, oder ob die Gestalt des Textes im Übersetzungsprozeß bewahrt werden kann. Da eine Vielzahl von Texteigenschaften zu berücksichtigen ist, deren intellektuelle Erhebung zudem sehr aufwendig wäre, ist die Anbindung von Textanalysewerkzeugen an TransRouter vorgesehen, welche die relevanten Textmerkmale automatisch bereitstellen.

Die Erhebung von Textmerkmalen durch Softwarewerkzeuge (z.B. Wortzählung) führt zu exakten, häufig numerischen Werten. In der Formulierung von Regeln wird der Experte hingegen Gebrauch von vagen Termen³ machen (*Lange oder komplexe Sätze/Texte sind weniger geeignet* für die maschinelle Übersetzung).

Die Erhebung der Textmerkmale ist selbst wiederum ein komplexer Vorgang, welcher der Optimierung bedarf. Die Erhebung der Textkomplexität oder der lexikalischen Abdeckung sind vergleichsweise aufwendig. Sie sollten nicht erfolgen, wenn etwa aus der gewünschten Sprachkombination bereits erschließbar ist, daß keine Übersetzungshilfen zur Verfügung stehen.

³ Dies ist noch durch Konsultation von Übersetzungsexperten zu bestätigen. Im positiven Fall sind Verfahren des Schlußfolgerens mit vagen Daten zu berücksichtigen (Spies 93).

2.3 Eigenschaften von Übersetzungsagenten

Als Agenten werden Personen oder Softwaresysteme erfaßt, die aktiv in den Übersetzungsprozeß eingreifen können. Ihr Beitrag zur Übersetzung ist einerseits von intrinsischen Faktoren abhängig und andererseits von der Qualität der einzusetzenden Ressourcen (Wörterbücher etc.), auf die im nächsten Abschnitt eingegangen wird.

Als Eigenschaften des Agenten sind Sprachpaare, Sprachkompetenz, der Übersetzungsaufwand, der in Bezug auf eine vorgegebene Übersetzungsqualität entsteht, sowie die optimal erreichbare Qualität zu modellieren. Darüber hinaus ist wichtig, welche Ressourcen von einem Agenten genutzt werden können und mit welchem Aufwand diese erstellt bzw. erweitert werden können.

2.4 Eigenschaften der verfügbaren Ressourcen

Als Ressourcen werden in diesem Zusammenhang passive Datenbestände bezeichnet, die von Agenten für die Übersetzung von Texten herangezogen werden können. In diese Rubrik fallen gespeicherte Übersetzungen, die von Translation-Memories in den Übersetzungsprozeß eingebracht werden, Terminologiebanken und Wörterbücher.

Die Qualität der verfügbaren Ressourcen, die in erster Linie ihren Umfang betrifft, wirkt sich auf die Übersetzungsqualität der eingesetzten Agenten sowie die entstehenden Kosten bzw. die Übersetzungszeit aus, wobei die Abhängigkeit von Ressourcen bei maschinellen Systemen größer ist als bei menschlichen Übersetzern. Eine Terminologiebank oder ein Translation-Memory erhöhen die Produktivität eines Übersetzers erheblich, ein maschinelles Übersetzungssystem ist ohne ein entsprechendes Wörterbuch nicht einsatzfähig.

2.5 Eigenschaften von Projektabwicklungsstrategien

Im vorangehenden Abschnitt wurde schon darauf hingewiesen, daß Übersetzungsprojekte durch den kooperativen Einsatz von Übersetzungsagenten, die jeweils ein Teilproblem des Gesamtprojekts lösen, durchgeführt werden. Ein sinnvoller Einsatz dieser Ressourcen ist also nur innerhalb eines umfassenden Arbeitsablaufs möglich.

Ein Arbeitsablauf bestimmt sich dabei über die Folge der Bearbeitungsschritte und die jeweils einzusetzenden Hilfsmittel und Ressourcen. Wurden unter dem Begriff der projektspezifischen Rahmenbedingungen die extern vorgegebenen Anforderungen an ein Projekt zusammengefaßt, so stellen die Abwicklungsstrategien die Maßnahmen dar, diesen Rahmenbedingungen gerecht zu werden. Es müssen die Workflows also auch über Attribute verfügen – wie Übersetzungsqualität, Kosten etc. Diese ergeben sich aus den Eigenschaften der eingesetzten Hilfsmittel und ihrem adäquaten kooperativen Einsatz. Sie müssen mit den Anforderungen des Projekts in Einklang gebracht werden. Dabei lassen sich folgende Phasen der Textbearbeitung mit den jeweils evtl. auszuführenden Operationen unterscheiden:

Vorbereitungsphase:

In diese Phase fallen alle Aktivitäten, die vor dem eigentlichen Übersetzungsvorgang anfallen:

Scannen

Anpassung von Formaten

Analyse der Texteigenschaften

Zerlegen des Texts: Umfangreiche Texte können aus Texten sehr unterschiedlicher Komplexität bestehen, an die differierende Ansprüche gestellt werden. In diesem Fall kann der Text zerlegt und mit den Verfahren bearbeitet werden, die für den jeweiligen Teiltext optimal sind. In einer einfachen Fassung dieses Arbeitsschritts besteht das kleinst-mögliche Segment der Bearbeitung in einer Datei aus dem vom Auftrag-geber übermittelten Datenbestand. Dateien mit äquivalenten Eigenschaften werden zusammengefaßt und auf die gleiche Weise übersetzt.

Aufbau von Wörterbüchern anhand der bei der Analyse bestimmten Defizite im Wortschatz.

Übersetzungsphase:

In diese Phase fallen die eigentlichen Übersetzungsaktivitäten, die für jeden der in der Vorbereitungsphase identifizierten Teiltexte gesondert durchgeführt werden.

Intellektuelle Übersetzung

Intellektuelle Übersetzung mit Unterstützung (Vokabular etc.)

Zugriff auf ein Translation-Memory

Maschinelle Übersetzung

Korrektur / Nachbereitung einer maschinellen Übersetzung

Nachbereitungsphase:

In der Nachbereitungsphase wird die endgültige Form des Textes erstellt.

Zusammenfügen der Teiltexte

Korrekturlesen

Formatierung

Speicherformat anpassen

3 Präzisierung der Projektziele und Skizzierung eines Entscheidungsmodells

Die begrenzte Laufzeit des Projekts TransRouter sowie die Notwendigkeit, wegen der erforderlichen Benutzerbeteiligung schnell zu einem funktionsfähigen Prototyp zu gelangen, legen eine Beschränkung der Fragestellung auf den Kern der für die Planung von Übersetzungsprojekten spezifischen Entscheidungsprobleme nahe. Somit werden zunächst alle Fragen der Ressourcenvergabe beiseite gelassen. Sie können von einem handelsüblichen Produktionsplanungssystem zufriedenstellend gelöst werden. TransRouter hingegen konzentriert sich auf den Auswahl angemessener Übersetzungsverfahren.

Auch dies könnte sich als ein sehr komplexes Planungsproblem herausstellen. Man denke an das folgende Szenario: Ein Text sei vom Deutschen in das Chinesische zu übersetzen. Für dieses Sprachpaar stünden weder Übersetzungssysteme noch kompetente Übersetzer zur Verfügung. Es findet sich die Lösung, den Text zunächst ins Englische und von dort ins Chinesische zu übertragen. Derartige Übersetzungswege sind zwar denkbar, aber deshalb nicht praxisrelevant, weil bei mehrfacher Übersetzung entweder eine angemessene Qualität des Zietextes nicht mehr zu erzielen ist oder untragbare Kosten entstehen. Es stellt sich heraus, daß es für die Praxis als ausreichend anzusehen ist, die in Abb. 2 skizzierten grundsätzlichen Vorgehensweisen zu unterscheiden (Konvertierung von Dateiformaten kann zusätzlich zwischen den Bearbeitungsphasen erforderlich werden).

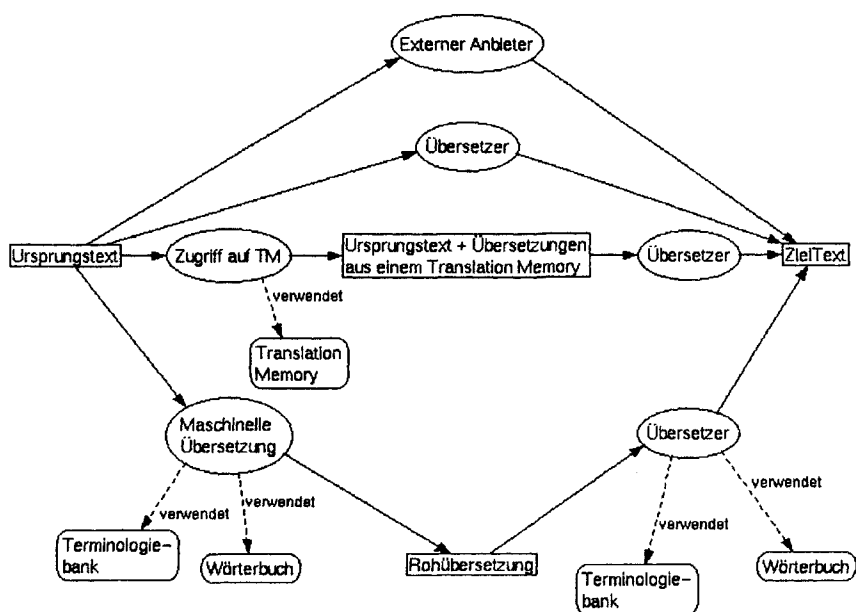


Abb. 2: Sinnvolle Bearbeitungswege eines Dokuments bei der Übersetzung (vereinfacht)

Es kann also folgender vergleichsweise einfache Algorithmus für den Entscheidungskern von TransRouter formuliert werden:

1 Finde für jeden Bearbeitungstyp Belegungen mit Übersetzungsagenten, die jeweils mit Ressourcen ausgestattet sind.

1.1 Suche für jeden Bearbeitungsschritt alle einsetzbaren Agenten. Teste dabei die einfach zu erhebende Vorbedingungen (z.B. Sprachpaar) zuerst.

1.2 Statte diese mit Ressourcen aus. Erzeuge Varianten der Ausstattung mit Ressourcen, die sich hinsichtlich Kosten-, Zeit, Qualitätskriterien unterscheiden. Offensichtlich ungünstigere Ressourcen (z.B. ein Translation-Memory, das eine geringer Abdeckung konkurrierende andere aufweist, in der Handhabung aber nicht effizienter ist) können unmittelbar ausgeschlossen werden.

1.3 Eliminiere alle Agenten / Ressourcenkombinationen, die gegenüber ihren Konkurrenten keinen Bearbeitungsvorteil aufweisen.

1.4 Aus den verbleibenden Varianten erzeuge alternative Instanzen der Bearbeitungstypen.

2 Bewerte diese Alternativen anhand der vorgegebenen Kosten-, Zeit- Qualitätsanforderungen.

4. Kosten / Nutzen

Die Einführung elektronischer Übersetzungshilfsmittel hat erhebliche Auswirkungen auf die Gestaltung des Arbeitsablaufs in Übersetzungsagenturen. Der Nutzen, der aus der neuen Technologie gezogen werden kann, differiert dabei erheblich je nach dem Typ des betroffenen Unternehmens. Während ein kleines Übersetzungsbüro, das auf hoch individuelle wissenschaftliche Texte aus einer bestimmten Domäne spezialisiert ist, vor allem von Fachwörterbüchern profitieren wird (weitere Betrachtungen über Übersetzungshilfen sind nicht erforderlich), so wird eine auf Lokalisierung von Software ausgerichtete Agentur insbesondere auf Translation-Memories oder maschinelle Übersetzungssysteme zurückgreifen können. Für das erstgenannte Unternehmen, das mit überschaubaren Mitteln operiert, ist geringerer Nutzen aus dem Einsatz eines Systems wie TransRouter zu erwarten. Er liegt nicht in der Ressourcenverwaltung sondern in der Unterstützung der Kostenkalkulation für Übersetzungsprojekte, die, sonst als Teil des Entscheidungsprozesses anzusehen, hier zum Selbstzweck wird. Für das zweite Unternehmen ergibt sich Nutzen aus der Auswahl der angemessenen Übersetzungsmethode, der Zuordnung von Ressourcen, die sonst häufig ungenutzt blieben, und damit in Vermeidung von unnötigen Arbeitsschritten. Natürlich wird auch hier die Kostenkalkulation relevant sein. Insgesamt kann der Beitrag von TransRouter zu einer erheblich verbesserten informationellen Absicherung der Unternehmensentscheidungen führen. Die entstehenden Mehrwerte (zur Theorie informationeller Mehrwerte vgl. auch Kuhlen 1995) sind in der besseren Integration bestehender Übersetzungstechnologie und ihrer effizienteren und effektiveren Nutzung zu

sehen. Sie erleichtern die strategische Entscheidung eines Unternehmens für neue Übersetzungstechnologien, die in ihren Implikationen für ein geordnetes Projektmanagement sonst noch nicht ausreichend eingeschätzt werden können.

Diesen Nutzen eines solchen Entscheidungsunterstützungssystems über die Zeit zu erhalten ist allerdings auch mit erheblichen Kosten verbunden, die aus der Aktualisierung der Datengrundlage für die Bewertung von Bearbeitungsalternativen resultieren:

Für alle relevanten maschinellen Übersetzungssysteme muß eine Matrix erstellt werden, die sprachlichen Problemfällen einen Kompetenzgrad des Systems zuordnet. Diese Matrix muß für neue Versionen des Systems aktualisiert werden. Es ist nicht davon auszugehen, daß Herstellerangaben in diesem Zusammenhang verläßlich sind. Es ist also eine unabhängige Evaluation der Systeme erforderlich, die z.B. auf der Basis von Testsuiten erfolgen kann, dennoch aber mit großem Aufwand verbunden ist.

Angaben über Produktivität von Übersetzungsagenten (z.B. der Aufwand zur Erstellung von Terminologien etc.) müssen aktualisiert werden, wenn Änderungen in der Benutzungsschnittstelle oder anderen relevanten Komponenten dieser Produkte erfolgen.

Kostentabellen externer Anbieter müssen ständig aktualisiert werden.

Erfolgt die Wartung des Datenbestandes nicht mit der nötigen Sorgfalt, so ist aus der Nutzung des Systems aufgrund irriger Bewertungen Schaden zu erwarten. Die einzigen Dienstleistungen, die vergleichsweise resistent gegenüber Wartungsproblemen sind, sind die Beurteilung von Translation-Memories und Vokabularen, die hier der Abdeckungsgrad vom System selbst festgestellt werden kann. Beschränkte sich der Nutzen von TransRouter dann auf denjenigen eines auf die Übersetzungsproblematik spezialisierten Informations-Repositories, so wäre für viele Agenturen immer noch einiges gewonnen.

Ausblick

Das Projekt TransRouter ist noch in einem relativ frühen Stadium. Ende 1998 wird ein erster Prototyp des Systems zur Verfügung stehen, der dann im Rahmen von Workshops den einschlägigen Nutzergruppen vorgestellt und von ihnen getestet werden soll. Federführend bei der Entwicklung des Prototypen ist das LRC. Die anderen Projektpartner sind für einzelne Softwarekomponenten, den Entwurf von Profilstrukturen und ihre Füllung mit entscheidungsrelevanten Daten zuständig. Die Tests des ersten Prototyps werden die Grundlage für den Entwurf weiterer Systemversionen bilden.

Literaturangaben

[Hutchins; Somers 92]

Hutchins, W. John; Somers, Harold L.: *An Introduction to Machine Translation*. Academic Press, London etc., 1992.

[Kuhlen 95]

Kuhlen, Rainer: *Informationsmarkt. Chancen und Risiken der Kommerzialisierung von Wissen*. Universitätsverlag Konstanz, Konstanz, 1995.

[Le-Hong et al. 92]

Le-Hong, Khai; Höge, Monika; Hohmann, Andrea: *User's Point of View of the Translator's Workbench*. In: *Translating and the Computer 14 – Quality Standards and the Implementation of Technology in Translation*. Aslib, London, 1992, S. 25-32.

[O'Sullivan 92]

O'Sullivan, Siegrun: *Quality Management in Translation*. In: *Translating and the Computer 14 – Quality Standards and the Implementation of Technology in Translation*. Aslib, London, 1992, S. 107-115.

[Spies 93]

Spies, Marcus: *Unsicheres Wissen*. Spektrum Akademischer Verlag, Heidelberg etc. 1993.

[Swanholm 92]

Swanholm, Flemming: „*The happy Triad*“ – *The Human, The MAT, and The MT*. In: *Translating and the Computer 14 – Quality Standards and the Implementation of Technology in Translation*. Aslib, London, 1992, S. 15-24.

[Turban 95]

Turban, Efraim: *Decision Support and Expert Systems*. Management Support Systems. Prentice Hall, 1995.

Information Management: Concept, Teaching, Applications

Leo Vodáček

Prague Economic University
Faculty of Informatics and Statistics
Department of System Analysis
130 67 Prague, W. Churchill Sq.4,
e-mail: VODACEK@VSE.CZ

Contents:

1. Concepts of Information Management
2. Concept of IM at Prague University of Economics
3. Teaching of IM
4. Applications of IM
5. Literature

Summary:

The paper explains the concept of information management as worked out and used by the Department of System Analysis of the Prague Economic University. Differences with some other typical interpretations of this term are exposed. Finally experience with teaching and practical applications of information management is briefly characterized.

1. Concepts of information management

In the theory as well as practice the concept of "information management" (IM) has not yet gained the clear and generally accepted interpretation. Three major reasons seems to be relevant for this situation:

a) Different **interpretations of the term "management"**. One can easily find tens of definitions, sometimes rather distinct from each other and expressed by eminent management authorities (Vodáček, Vodáčková, 1996).

In the further explanation we do understand under the term of management the set of proven approaches, methods, experience and recommendations used by executives (managers) for coping with specific activities to achieve the goals of the organization. These specific activities are sequential managerial functions, like planning, organizing, staffing, leading and controlling, being the core of managerial activities.

b) Different **interpretations of the term "information"**, as we know them from the vast world literature (for example Vodáček, Rosický, 1997).

In the further explanation we accept that "Information is data endowed with relevance and purpose. Converting data into information thus requires knowledge" (Drucker, 1990, p. 202).

c) The **changing leading ideas about the major use and roles of IM** since 1960s to the present days.

According to our opinion one can follow the development of the term and concept of IM in three principal stages.

In the **first stage** (the second half of 1960s) the role of IM had been primarily identified with skill of the right choice and use of data, methods and approaches for ensuring "engineering efficiency" in solving technical and technological tasks. Probably one of the first use of the term IM was by R.S. Taylor and his colleagues during the conference to engineering efficiency in 1966 (Taylor, 1966).

Later in the 1970s the term IM had begun to be frequently used for some tasks of data processing, for example in providing order into information resources in technological documentation, libraries etc.

In the **second stage** (the end of 1970s and during 1980s) the term IM had been used primarily for methods and approaches of professional working in the field of informatics. The accent had been on the managerial approaches and techniques for efficient handling of information resources, primarily the efficient design, implementation and use of information systems (IS). Efficient information processing through IS had been usually considered as final goal of IM.

For example Michael J. Earl suggests that "information management comprises planning, organization and control of information resourcesEffective information management requires planning methods, control procedures and organizational arrangements to be congruent with each other. In the main they also must fit the management practices of the firm at large" (Earl, 1989, p.24).

Due to the fact that planning, organization and control constitute key parts in most concepts of sequential managerial functions - one can see, that in Earl's view IM is considered as managerial support for dealing with information.

In the second stage in the development of IM concept one can already see and appreciate the emphasis on transdisciplinary integration of information

processes with management approaches. It helped to respect economic criteria of efficiency in designing, implementing and using IS.

On the other hand one could frequently observe absence or insufficient participation of the end users of information resources or IS in their creation. They rarely posed questions and looked for answers how IS/IT could help significant innovation of managerial work (to formulate new managerial tasks or philosophy of the solution of traditional ones).

One can say in a bit simplified way that IM was considered as a set of managerial methods and approaches serving to the needs of technologically oriented informatics (IS/IT). Or "doing things right" (efficiency). It has been reflected in the growing introductions of Information Managers, Chief Information Officers etc. in organizations using IS/IT.

In the **third stage** (since the beginning of the 1990s) the concept of IM reflects more and more the managerial priority for the use of IS/IT. The accent is on the final use of IS/IT, i.e. innovative solution of managerial tasks, primarily the effective fulfillment of the mission and goals of the organization. One can see the growing attention being paid to the right effectiveness of the information processing ("doing the right things").

As the first example we can quote the "Association of Information and Image Management" (AIIM in Great Britain) defines IM as "the effective production, storage, retrieval and dissemination of information in any format and on any medium to support business objectives" (Best, 1996, p.4).

As the second example we can quote William J. Martin's definition of IM as "management of information resources of an organization in pursuit of its aims and objectives. As such it requires the application of standard management processes of planning and control, while seeking to ensure the day-to-day flows of information for decision-making and the concurrence of information and business strategies within the organization" (Martin, p. 171).

In the third stage in the development of IM concept there is a growing stress on the final effect for the use of IS/IT, i.e. on the "value matrix" of managerial work. The priority of managerial goals and objectives came into forward being reflected in the requirement of effectiveness for the use of IM too ("doing the right things"). And effectiveness is not put into contradiction, but in coalition, with the dual additional requirement for efficiency ("doing the things right").

2. Concept of IM at Prague University of Economics

Since the beginning of the 1990s the Department of System Analysis of the Prague University of Economics was working on the concept of IM too. It was caused not only by needs of the modern pedagogical process corresponding to the standards of advanced industrial countries, but at the same time by needs of a very rapid dissemination of modern management in Czech microeconomics sphere. Newly privatized firms and other organization asked for our

participation in the rethinking and redesign of their managerial systems, incl. IS/IT. At the same time they required managerial simple and for industrial practice understandable methodological recommendations.

Confronting our positions and needs with several concepts of IM in the world - we soon declared our adherence to the managerial priority for the use of IM (correspondence to the "third stage of development"). Further the need of system approach was stressed - as a substantial help for the right formulation and solution of managerial tasks. In this sense **under the term of IM we understand the appropriate use of modern means and methods of information technology and system approaches for coping with information needs for managerial activities.** Effective as well as efficient information processing are considered as important requirements posed on applications of IM.

In our concept **IM represents the transdisciplinary intersection of three principal components. They are: management, informatics and system approaches (system analysis and synthesis).** Due to character of specific applications knowledge from some other additional components may be integrated, for example from economy, marketing, financing, pricing, psychology, sociology, operation analysis, politology, etc.

For its use is **IM considered as multifunctional.** The significance of individual components is not considered only from the position of the their partial professional contributions, but in the new synergetic quality of a whole IM given by their goal-oriented integration.

By being brought into a new relation with one another, management, informatics and system approaches have undergone varying degrees of modification for their creative use. It enables sometimes new formulations of managerial tasks, new ways of their solution and interpretation of results (for example integration effects trough different types of coordination in organizational structures, process integration trough BPR, effects of structural mergings).

Due to the extend of our paper, the more comprehensive insight into individual components and the ways of their synthesis will remain outside further explanation, but was already published (Vodáček, Rosicky, 1997). Informatics as the key innovating agent in modern management plays the leading role for the effective and efficient performance of managerial functions, roles or "critical factors". Attention is paid to the growing impact of information processing in transforming managerial theories as well their practical applications. The new managerial environment is being reflected, for example on the background of P. Drucker's ideas about "knowledge based society" (Drucker, 1990), "learning organization" (Senge, 1990), "organization learning" (Argyris, 1992), MIT Study about the corporation of the 1990's (Scott Morton, 1991) or "intellectual capital" (Edvinson, Malone, 1997).

In solving practical tasks we do not use only classical structural arrangements corresponding to frameworks of managerial functions, but with growing

frequency the up to date systemic concepts defining "critical factors" of particular problems.

3. Teaching of IM

Our concept of IM was in principle worked out in 1992 and gradually in following years partially innovated, tested and applied for the solution of several managerial tasks, particularly on the microeconomics level. At the same time corresponding courses were introduced at our Prague Economic University and later on some other Czech Universities too.

The starting course of "Introduction to Information Management" (in the 3. year of study) provides students with understanding of principles of IM. It stresses the role of IM, its transdisciplinary character and systemic approach to the integrative use of applied disciplines.

In further two years parallel to the detailed course of "Information Management" there are **obligatory courses** of

- Management Information Systems,
- Methodology and Analysis of Systems,
- Decision Support Systems,
- Management of Projects and their Economy,
- Innovation and Design of Systems,
- System Analysis (Seminary 1 and 2).

Corresponding to their interest and needs students have **elective courses** too, for example

- Information as an Economic Resource,
- Natural Thinking and Artificial Intelligence,
- Managerial Communication,
- Theory of Information,
- Design of IS/IT.

The names of the courses mentioned above present a rough idea about the contents and structure of the teaching process. For providing reasonable flexibility and respect of individual preferences of students the credit system has been applied.

4. Applications of IM

The concept of IM as worked out by our Department of System Analysis has helped us substantially in working out the set of systemic recommendations for the solution of different acute and complex managerial tasks. They were used not only for pedagogical purposes in our courses, but primarily for consulting activities of our department in collaboration with Czech macro- as well as micro-economic practice.

Typical tasks being solved with the help of IM methods were problems corresponding to rather complex socio-economic systems (the higher half of Boulding's taxonomy levels). They included

- factory planning systems based on IS/IT,
- design of integrated production management systems,
- design of strategic alliances and other forms of integration of Czech and foreign firms,
- applications of business process reengineering (BPR),
- strategic restructuring of Czech industrial firms.

In order to illustrate in more detail at least one application of IM we would like to mention our methodological approach to the task of **strategic restructuring in Czech industrial firms**. The first reason for selecting this problem is, that strategic restructuring can be considered at the present time as the key task of our microeconomics transformation process. At the same time - and it is the second reason for our choice - the example provides good opportunity for the presentation of the typical spectrum of methods and approaches being used in IM (modeling, application of hard- and soft- approaches, application of modern management thought stressing the new roles of IS/IT, etc.).

Strategic restructuring has been defined as the **qualitative transformation (innovation) of inputs, outputs and structures of an industrial firm in order to meet the goals of its business**, and to cope with this task in compliance with economic requirements of **effectiveness and efficiency**.

In difference to many "practice-oriented" recommendations existing in the Czech Republic our "integrated approach" strictly differentiates the contents and goals of strategic restructuring, from partial restructuring processes. It is known, that the present restructuring effort of most Czech industrial firms has been wrongly focused only on financial restructuring or/and organizational restructuring. This has led to non-systemic solutions and caused further economic difficulties in factories, as well as the Czech economy as a whole.

The set of structural arrangements being appropriate for the analysis and synthesis of restructuring process has been identified as

- production program structure,

- production base structure (production structure),
- information structure,
- functional structure,
- organizational structure,
- personal (professional and qualification) structure,
- financial structure.

Individual structures are considered as mutually interlinked, have mutual positive or negative impact (intercorrelations). In the restructuring project and during the restructuring process the methods of IM for their harmonization are used. At the same time these fundamental structures might be extended by some specific structures, like R&D structure, innovation structure etc. Functioning of all structures is to a certain, but different degree affected by their internal and external environment (social, economic, technological, market, legal, cultural and even other "impact factors").

Making analogy to concepts of "critical success factors" (Vodáček, Vodáčková, 1996) or "components of organizations" (Scott Morton, 1991) **we have proposed for our Czech firms 4 mutually interlinked factors, which we consider as primary important for the process of strategic restructuring** (strategic concepts, projects). They are

- **business activities firm is and will be engaged in** (including primarily the production program structure and production base structure),
- **organizational methods, structures and processes** creating rational order and forms for the functioning of a firm (including primarily the functional, information and organizational structures and processes),
- **methods and systems of information processing**, as instruments for coping with the requirements of selected critical factors (including primarily the activation of structural elements and their ties in the systemic whole),
- **human factor**, as the integrative element for the effective and efficient realization of previous three tasks (key role of personal structures with their "soft-characteristics").

In order to be easily understandable and acceptable to the majority of Czech top and middle managers the choice and formulation of factors has preferred simplicity. At the same time the necessary integrated processing of these factors can be easily interpreted as the **unification of content, forms and instruments of managerial activities with the key role of people**. Factors and work with them was formulated in the set of understandable recommendations corresponding to the modern IM progress and corrected with regard to the feasibility in conditions of Czech firms (alternatives). It has been described in detail (Vodáček, 1998).

The clear system arrangement of the restructuring process provides the good starting point for the qualitative innovation of former business activities. It

facilitates modeling, integrated redesign and integrated management of core and supporting business processes through BPR methods, as well as clear statement of their value metrics (Johansson, 1993).

The application of IM (primarily due to IS/IT and BPR) in restructuring effort of business processes allows to use the information simultaneously in as many places as it is needed. With the help of computers and telecommunication networks it allows to have the same information at the same time in as yet locally separated organizational units. It makes possible to use methods of management by objectives (MBO) and to participate on the solution of local as well as global problems. In other words it facilitates to combine benefits of centralization and decentralization - one of constant challenges of Czech industrial practice.

5. Literature

Argyris Ch.: On Organizational Learning. Blackwell Publishers, Cambridge 1992

Best D.P.: The Fourth Resource - Information and its Management. Aslib/Gower, Aldershot 1996

Davenport T.H.: Process Innovation: Reengineering Work Through Information Technology. Harvard Business School Press, Cambridge 1993

Davidow W.H., Malone M.S.: The Virtual Corporation - Lessons from the World's Most Advanced Companies. Harper Collins Publisher, New York 1992

Davies L., Ledington P.: Information in Action - Soft System Methodology. Macmillan, Houndmills 1991

Drucker P.F.: The New Realities. Mandarin Paperbacks - Michelin House, London 1990

Earl M.J.: Management Strategies for Information Technology. Prentice Hall, University Press Cambridge 1989

Edvinson L., Malone M.S.: Intellectual Capital - Realizing your company's true value by finding its hidden roots. Harper Collins Publ., New York 1997

Galliers R.(ed.): Information Systems Research - Issues, Methods and Practical Guidelines. Blackwell Scientific Publications, Oxford 1992

Hamel G., Prahalad C.K.: Competing for the Future. Harvard Business School Press, Boston 1994

Hammer M., Champy J.: Reengineering the Corporation - A Manifesto for the Business Revolution. Harper Business, New York 1993

Hammer M., Stanton S.A.: The Reengineering Revolution - A Handbook, Harper Collins, New York 1995

Johansson J.H., McHugh P., Pendlebury A.J., Wheeler W.A.: Business Process Reengineering. John Wiley, Chichester 1993

Martin W.J.: The Global Information Society. Aslib Gover, Aldershot 1995

Peters T.J.: Thriving on Chaos - Handbook for a Management Revolution. Macmillan, New York 1987

Savage C.M.: Fifth Generation Management - Co-creating Through Virtual Enterprising, Dynamic Teaming, and Knowledge Networking. Butterworth - Heinemann, Boston 1996

Scott Morton M.S.(ed.): Corporation of the 1990s - Information Technology and Organizational Transformation. Oxford University Press, New York 1991

Senge P.M.: The Fifth Discipline: The Art and Practice of the Learning Organization. Doubleday Currency, New York 1990

Taylor R.S.(ed.): Information Management in Engineering Education. Bethlehem PA, Lehigh University Press 1966

Vodáček L., Vodáčková O.: Management - Theory and Practices for the 1990s. Management Press, Praha 1996 (in Czech)

Vodáček L., Rosický A.: Information Management - Concept, Role and Applications. Management Press, Praha 1997 (in Czech)

Vodáček L.: Tasks for Restructuring of Czech Companies. Politická ekonomie č.5/1998 (in Czech)

Watson G.H.: Business Systems Engineering - Managing Breakthrough Changes for Productivity and Profit. John Wiley, New York 1994

Wilson B.: Systems - Concepts, Methodologies and Applications. John Wiley, Chichester 1991.

SO-grams: a personal visualisation toolkit for intranet users

Elisabeth Davenport, Kathy Buckner

**Department of Business Information Management,
Napier Business School, Napier University, Sighthill Campus, Sighthill Court,
Edinburgh EH11 4BN
e.davenport@napier.ac.uk**

**Department of Communication and Information Studies,
Queen Margaret College, Clerwood Terrace,
Edinburgh EH12 8TS
k.buckner@mail.qmced.ac.uk**

Contents

Abstract

Background

The origins of the social browser concept

Level one: 'draw and describe'

Level two: 'draw and describe' and 'point and click'

Level three: mobilising SO-gram data

Field conversations

Social browsing in local government

Social browsing in the food and beverage sector

Social browsing in a global computer manufacturer

Conclusion

Acknowledgements

References

Abstract

The problem: visibility in the virtual workplace

We are interested in the problem of how to share human expertise in large distributed organisations whose work is carried out remotely by means of intranets, or comparable media platforms. In such environments, expertise cannot be transferred in the face to face exchanges or direct observations that physical proximity affords. Knowledge management in the virtual workplace must thus include an added dimension: the knowledge of each other's skills, experience and expertise, which is taken for granted (or tacitly acknowledged) in interactions in physical space. In the text which follows, we present a generic interface for social browsing (the 'SO-gram') which may improve 'visibility' (or employee's awareness of each other) in the virtual workplace. We describe the origins of the browser in fieldwork in a real-world community, and discuss its viability as a tool for social navigation in the context of three different organisations: local government, an independent medium size enterprise in the food and beverage sector, and a global computer manufacturer.

Background

In identifying the problem of community myopia, and in developing a 'social browsing' tool which might address the problem, we have drawn on different literature sets. The first is social network analysis, and is concerned primarily with the charting and analysis of social cohesion and the graphical display of cognate groups that are key components of the study of group dynamics (Grosser 1991). Haythornthwaite's work (Haythornthwaite 1996) on social networks as a tool for the study of information exchange, and recent conference presentations by Huotari (Huotari 1998) and Sonnenwald (Sonnenwald 1998) have proved to be particularly useful in our study, as has work on family support (Tracey and Whittaker 1990), where reliable representations of social space result in therapeutic interventions. Though such work demonstrates the explanatory power of social networks in terms of data analysis and data visualisation, they do not, as far as we are aware, address the issue of how social network data might trigger interaction, or, in the terms of our study, improve social visibility.

Another important source was the literature on navigation tools. The 'trigger' effect is covered in the literature on navigation, most pertinently in discussion of 'push technologies' like *Firefly* (Lyons 1998), which puts readers/users in touch with like-minded users or like-material on the basis of experiential profiles. Although *Firefly* engines have been adopted in some sectors of e-commerce (Amazon.com) we have not found references to this type of technology as a social trigger within organisations. As work by Vishik (Vishik 1997) and Wexelblat and Maes (Wexelblat and Maes 1998) indicates, the focus in the literature on organisational intranets is on navigating texts or documentation. With the notable exception of Dieberger (Dieberger 1997, Dieberger 1998) who discusses 'social navigation', those researching the virtual workplace do not appear to treat human expertise as a navigable resource. Dieberger's work

provides an introduction to the webspaces and interactive genres that engage humans in virtual environments, but offers little guidance on how interaction or synergy might be engineered.

We are also aware of studies of the management and legitimisation of intranets by Rosenbaum (Rosenbaum 1997), and Sheepers and Damsgaard (Sheepers and Damsgaard 1997), and of workflow architectures for the virtual workplace (Ames *et al* 1997). Though these may well describe developments when partners are familiar to each other, they do not address the problem of initial contact. As we indicate above, our proposed networked version of the SO-gram, may both enhance observation of interactions, and overcome organisational myopia to promote interactions. In developing the concept, and designing prototypes, we have drawn heavily on the work of Spence (Spence 1998) and his colleagues on browsing. Browsing, in Spence's terms, is a general scoping activity which allows members of a community to find points of interest, and pursue them within the constraints of the effort required. Criteria for an acceptable social browser in a given context must take account of a group's hedonic calculus. We discuss these issues, albeit at a simple level, in the text which follows.

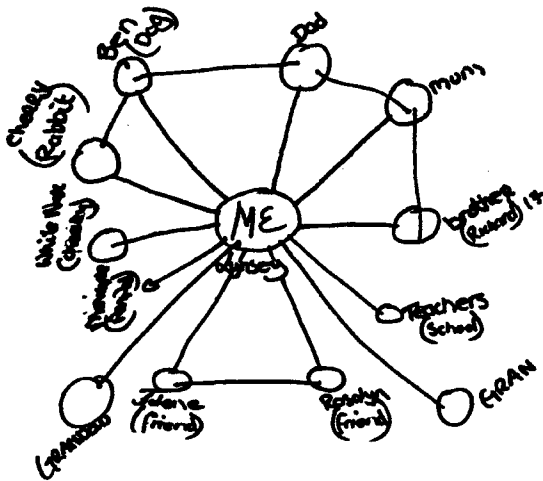
The origins of the social browser concept

For the past year, we have collaborated on a project which seeks to investigate the application of information technology to the enhancement of social cohesion within a small geographic community in the West End of Edinburgh. Our dialogues in the locality, with individual residents and members of community groups, have focused on group interaction. Under this rubric, we have tried to elicit who is significant in our respondents' lives, and how that significance is manifest and articulated in exchanges between respondents and their significant others. We have attempted to discover how respondents rank their relationships, and whether patterns of behaviour reflect rankings in a way that allows zones of comparative significance to be identified and described in terms of attributes. In addition, we have explored the viability of a tool for extending relationships beyond the immediate circle of friends and friends of friends. As a result of our fieldwork we have identified a generic problem 'community myopia' (Davenport *et al* 1998a) which may be overcome with tools that enhance 'community visibility'.

Level One: 'draw and describe'

To track the relationships of participants in the community project, we developed a simple conversational prop, the 'SO-gram', or 'significant other' gram (Davenport *et al* 1998b). Each participant was asked to draw a simple node and arc diagram of important others (people and animals) in her or his life, and to place himself or herself at the centre of the 'map' (Fig. 1). The nodes were drawn in different sizes to reflect the intensity of the relationship - a large node is more 'significant' than a small node. Participants were then asked to talk about what 'significance' implied (fondness, trust, frequency of contact, long-termism and so on), and to describe the relationship (represented by the

Figure 1. Level 1 SO-gram

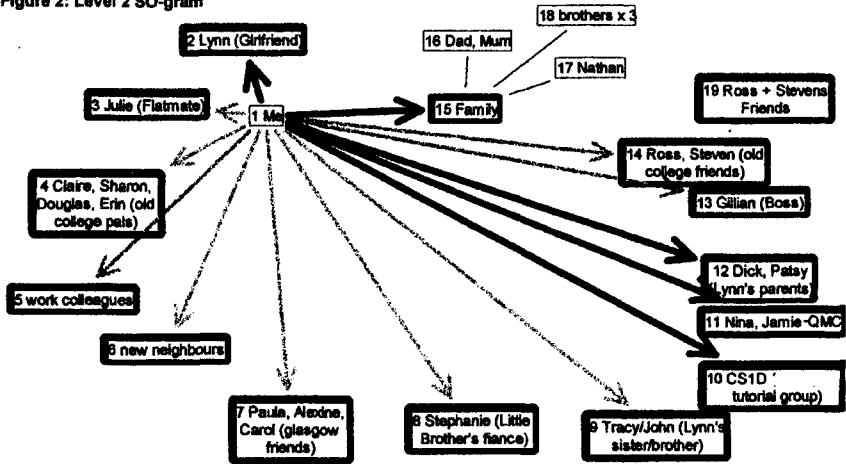


arc) between the 'me' node and its neighbours. Though useful as a conversational prop (participants found the diagram easy to draw, and to talk through), the paper SO-gram has two obvious limitations as a social browsing tool: firstly, it represents only one-to-many relationships, and is thus unlikely to take an individual outside an existing circle; and it is navigated in a simple chaining process, which is both time consuming, and unwieldy where 'multiple' paper SO-grams are stacked.

Level Two: 'draw and describe' and 'point and click'

In Level Two, we used simple computer tools to model interactions, using SO-gram data based on a version of the basic diagram which was modified by asking participants to follow standardised procedures for annotation of the node and arc diagram, and to use prescribed criteria to 'zone' their interactions. Not only did these constraints allow the SO-gram data to be transferred into two off-the-shelf application packages, but they were also appreciated by participants, who appeared to take less time to offer a more consolidated view of their social interactions. The two packages, *Decision Explorer* (Decision Explorer 1998) and *Inspiration* (Inspiration 1998), have been designed to help in the representation of qualitative information in complex environments, by allowing links between different concepts to be established and manipulated, and facilitating the analysis of concept maps (Fig. 2).

Figure 2: Level 2 SO-gram



There are several advantages in using a computer tool. Both tools allow nodes to be moved around in the SO-gram, so that individuals from the same groups can be positioned near to each other. *Inspiration* allows free form circles to be drawn around the group to allow visualisation of zones. Both tools allow users to zoom into the SO-grams of neighbours by entering a given node, and both packages allow users to store data for future analysis on the types of exchanges which are taking place. Using import/export functions it is possible for data from qualitative data analysis packages to be linked to concept mapping tools. Instead of concepts, we used people as node labels.

Though the simple computer based tool clearly increases the analytic scope of the SO-gram as a social browser, we identified several drawbacks. In terms of navigation, *Decision Explorer* and *Inspiration* tie the user to a 'chaining' process which may be inadequate in the context of a complex organisation, and we suspect that their optimal capacity may be limited. In addition, the resource implications may be high where the tool is implemented and maintained by small groups on low budgets, and some users may find the computer tool too demanding unless a facilitator is present. Computer-based SO-grams may be modified in ways which make them easier to construct, and improve their efficacy as carriers of focused information. We are currently 'tidying' the protocol for constructing SO-grams to derive a series of versions which can accommodate different participant levels of comfort.

Level three: mobilising SO-gram data

The tool has potential as a probe to explore the underlying communication structure in larger groups, where it might provide a mechanism for collecting parsimonious data for personal profiling. If, for example, SO-grams were to inform the design of an agent-based message system for a given population, they would need to deliver information on:

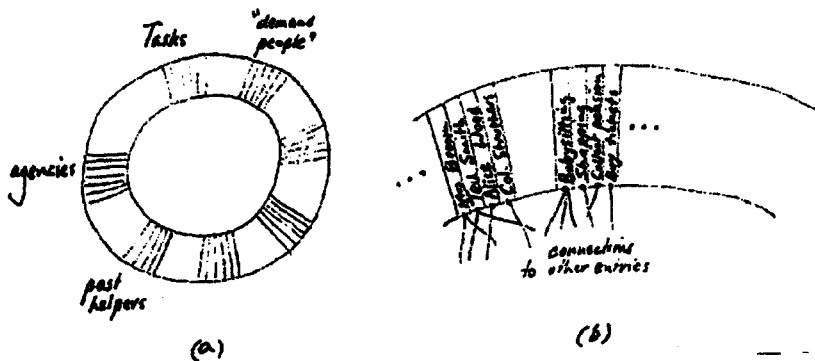
- motivation for communication;
- type of exchange;
- initiation-response patterns;
- sites for interaction.

Between June and September 1998, we worked with a commercial network mapping tool, *Netmap*[®] (Netmap 1998) to test the usefulness of SO-gram data embedded in a complex computer application (Connolly 1998). We derived certain 'mobilisation' criteria for the application to function as a 'rich' social browser which would overcome the drawbacks identified in our application of the simple computer-based version:

- the tool must be capable of representing each entry in the SO-gram as a different node, and must allow these to be labelled accordingly - it must support differentiation;
- the tool must be capable of representing the links between nodes and differentiating between links of different types, displaying only links of a certain type if required - it must support extraction;
- the tool must be capable of simplifying the connection networks when the number of connections is so high as to detract from the clarity of the information presented - it must support filtering;
- the tool must be capable of accessing information inside nodes - it must support data mining.

Netmap[®] was originally developed on the UNIX platform and its graphical user interface is based on the X-windows system, It is now available for the Windows NT platform; however, the GUI still requires X-windows to be installed on NT machines. The software is licensed and a key code can be obtained from the developers. The installation process is relatively straightforward but the assistance of a systems administrator may be required when installing into a networked environment. The package accommodates complicated manipulations of the sort which we wish to represent, and is thus pertinent to our search for a social browser that overcomes community myopia in a range of contexts.

The features of a basic *Netmap*[®] are presented in Fig. 3. Nodes are presented as small sectors around a circle, each with its own label. These sectors are collected into groups which can be differentiated by colour, as can the links between members of different groups within the main circle. Beside each link is a smaller 'satellite' circle made up of all the nodes of that group, within which a user can find all the links within that group. The software can select links of a designated type and also users to extract information using threshold values as a filter.



(Fig. 3)

An analyst who wishes to improve his or her ability to observe social interactions in a community can choose a 'discrete network' grouping strategy that allows small isolated groups to be identified and displayed. In addition, an emergent grouping strategy allows an analyst to see the weaker connections between groups without combining them into a single supergroup on the basis of one or two transgroup interconnections. This feature would be useful for managers wished to evaluate the effects of strategies for social cohesion. Netmap® triggers social interaction by allowing users to search on individual nodes (a task or a location, for example), to find out who has previously linked to whom on these, and to mine for further details on the flipcharts that lie behind each node. These features would support a member of an organisation wishing to identify an experienced (and trustworthy) partner for a project.

Field conversations

As we indicate above, we see scope for 'social browsers' in the virtual workplace, or intranet-driven organisational work. Where the browser allows individuals, and groups, to observe 'visualised' interaction, or where it triggers interaction outside the scope of observable activity, it may contribute to organisational synergy and social learning.

From May to August 1998, we held a series of interviews with representatives from three organisations in an attempt to scope the viability of the browser, and identify contexts where more sustained research might be carried out. Organisation A is an information and library service in a local government authority which has recently been re-structured as an independent body. Organisation B is an independent medium sized enterprise in the food and beverage sector which has experienced a series of major transformations in the past three years: management buy-out, flotation, and, automation of the sales force. Organisation C is the Scottish arm of a major global computer manufacturer. The interviews focused on two dimensions of the social browser: its potential as an observatory, and, its potential as a trigger for social

interactions that might otherwise not have happened.

Social browsing in local government

DM is the director of one of the smaller departments in Organisation A, which, as we indicate above, has recently undergone major restructuring and has been split off from a larger authority, to be run as an autonomous unit. Overall, DM is responsible for 55 people, scattered across 11 branch offices, and two mobile service units. In addition to daily running responsibilities, DM is increasingly involved in the preparation of bids for funding, as the local authority budget is tight. These bids cover all aspects of his work - from capital funds for building through the Private Finance Initiative, funds for the re-skilling of personnel to fit them for the internet age, to EC funds which will allow DM's department to acquire much needed equipment. DM summarises the three main areas where SO-grams might be useful as:

- to assist him, and others, in their understanding of the newly configured (and constantly fluctuating) organisational structure and in their ability to contact appropriate colleagues for given tasks;
- to assist him in his search for funds, both in terms of locating funders and partners where funders require collaboration as a condition of grant etc;
- to act as a personal organiser, and a framework within which to evaluate career trajectories.

DM can see that levels one, two and three might all be useful in different ways, depending on the status of the individual who uses them. Level one is most appropriate for the 'personal organising' function, and level two for the 'dynamic organigram'. As overall manager of the unit, however, DM would find level three SO-grams useful in identifying areas where better connections might be made among his staff, or identifying cases where staff are not engaging with a particular task or a particular site. He would also welcome a system that could put him in touch with potential partners, on the basis of others' personal experience; much time can be wasted meeting with partners who are not suitable.

Though he can see many benefits rising from the application of SO-grams, DM is sceptical that the local authority in its present state of IT development might adopt them as an organisational standard, as the benefits from levels two and three depend on mutual connectivity, and the local authority is reluctant to provide this for reasons of security. He would be happy to implement 'level - one' prototypes within his own unit if the cost was controlled. One way to implement mutually networked SO-grams might be to develop 'social browsers' as a sectoral standard for collaborative interfaces within library and information services. He would be happy to be included in any grant proposals to further this.

Social browsing in the food and beverage sector

MS is a senior administrator in a medium sized enterprise (around 500

employees) in the food and beverage sector, which, as we indicate above, has undergone major re-structuring since 1993. A management buy-out was followed by flotation on the stock market in 1996. A 'level one' SO-gram drawn by MS clarified six main areas of responsibility: Personal Assistant to the Managing Director, induction and training, coordinator of the Salesforce, stationery and office supplies, office documentation. (MS indicated that drawing the SO-gram in itself had proved useful). Two further emerging roles are assistant to the IT manager, and assistant coordinator of financial services.

In most of these roles, MS acts as a gatekeeper, putting people in touch with those they need to work with. She sees her work with office documentation as part of this: she was solely responsible for designing and introducing standard forms and protocols for all office procedures throughout the enterprise, which she sees as a way of making sure that people have access to the same information and can share it effectively with the minimum of effort. This groundwork has prepared the way for computerisation of the company, which is rapidly adopting an enterprise wide computing strategy in response to the demands of distributors and suppliers, and to pressure within the sector. MS could see scope for 'level two' SO-grams in five main areas:

- as personal visualisation tools for daily planning and career development;
- as a general interface for presenting the company organigram and job descriptions with contact points;
- as interfaces which would enhance the induction process (whenever an employee is put in the position of having to 'start over' in relation to a new project, for example) by offering point and click access to colleagues at the level of the organisation, the project and so on;
- in relation to the expanding global sales initiatives, as a means of making the work of regional reps. visible to the rest of the company,

and,

- as a key component of the 'retrack' system. The field sales force have recently been equipped with laptops and mobile datalinks, which will allow them to retrieve and input information on the fly. A SO-gram would provide rapid access to their contact networks, and would allow new clients to be added. 'Significance' in this case might be measured in terms of the 'temperature' of the contact - cold, warm hot etc.

MS did not feel able to comment on 'level three' SO-grams, as she felt that that kind of analysis would be done by upper management; in this context it might be of use in the annual planning sessions, when next year's targets are set in response to market trends, and five year plans are drawn up.

MS was very positive about level one and level two SO-grams, and would like to see them built in to the emerging computing platform in Company B. The main problem might be time consuming data input, but many of the problems

associated with that have been dealt with in the office document standardisation project. At present, interface design and the social aspects of computing are secondary issues in the company, as accounting has dominated IT initiatives up till now. MS does not doubt that pressures from the trade will, however, make the 'human connection' factor increasingly important. Two areas where she would like to make a start with SO-grams in Company B are:

- personal evaluation and induction;
- the laptop initiative with the salesforce, where they could be combined with the communications facility.

Social browsing in a global computer manufacturer

IR is a senior development engineer in a global computer manufacturing firm. His immediate workgroup (located at the same site) consists of 10 - 15 colleagues depending on the project to hand. In addition, he has contacts with the Marketing and Manufacturing departments, with colleagues on other projects, and with colleagues on the sister site in the US. His social networks are to some extent dictated by the projects that he works on, most of which have a time frame of eighteen months to two years. Strong bonds develop between member of the project teams, who rapidly become 'significant' to each other.

The group currently uses an intranet for communication and sources, and has web access; one of the group members acts as a gatekeeper, particularly for online information. As existing systems already provide good networking facilities, IR thinks that level three SO-grams would be of most interest to his group because of the added value that they provide in terms of triggering social/collaborative links that might not otherwise have been made. For example, a newcomer to the firm, added to a team with a specialist function, might need to locate others in the company with expertise in that function who had worked successfully with other team members - a level three SO-gram might facilitate the connection.

IR is aware that in other computer companies, with a less hierarchical structure than his own, teams work to shorter time frames, and place more responsibility on new-comers to organise and configure their own groups. In this case, the need to find 'unknown' talent may be greater, as team members cannot rely on the channels that exist in established company divisions. In companies like this, an individual's work load may be more heterogeneous than in IR's firm, and a 'tracking' interface would obviously be of use.

In terms of implementation, any SO-gram browser must be unobtrusive, and require no greater effort than existing systems. A level three system would require an administrator; this might be added to the existing gatekeeper's portfolio. In addition, if the browser proved to be successful, it would be likely that the company would wish to develop it themselves as an in-house or commercial product.

Conclusion

We have outlined a problem, 'community myopia', that can affect group members who may wish to accomplish tasks together, but are unaware of each other, and have described a possible prosthesis - a 'social browser' to improve corporate visibility. We suggest that both the problem, and the proposed solution, may be relevant in the virtual workplace, where members of organisations cannot enjoy the easy contact afforded by physical proximity. Field 'conversations' (discussions of concept scenarios in three different organisations) revealed that:

- social browsers are best presented to organisations as systems that operate at different levels;
- the appropriateness of the level used will be determined by the state of development of technology (infrastructure and workforce experience) in the organisation;
- the amount of effort that an organisation is prepared to invest in a social browser will be related to the perceived added value offered by the tool.

Acknowledgements

The work described in the early part of the paper was supported under the Living Memory Project (LiMe) (Esprit 13/25261); later work was supported by funding from Queen Margaret College. In addition to acknowledging our funders, we would like to thank the following: Professor R Spence for sharing ideas on social browsing; Kirsty Barr for her work on SO-grams; Angus Whyte for his work with *Decision Explorer* and *Inspiration*; Justin Robertson and Reuben Connolly for their work on the Netmap® project; and, the three organisational representatives whom we interviewed.

References

- [Ames *et al* 1997]
Ames, C.K, Burleigh, S. and Mitchell, S.J. *A web based enterprise workflow system*. In Group 97. New York: ACM, (1997) pp.214-220.
- [Connolly 1998]
Connolly, R. *Netmap® as a visualisation tool for social interaction*. Information Engineering Section Report. Department of Electrical and Electronic Engineering. Imperial College of Science, Technology, and Medicine: London, September 1998.
- [Davenport *et al* 1998a]
Davenport, F., Buckner, K. and Spence, R. *Making interactions visible: tools for social browsing*. Presentation at the workshop, Le Travail Humain, Paris 14 September 1998.
- [Davenport *et al* 1998b]
Davenport, E., Buckner, K. and Barr, K. *SO-grams: work-in progress on a simple conversational prop for navigating social space*. In Hook, K., Munro, A. and Benyon, D (Eds). Workshop on personalised and social navigation in information space, 16- 17 March 1998. Swedish Institute of Computer Science, SICs Technical Report T98.2. Kista, Sweden, (March 1998), pp.45-54.

[Decision Explorer 1998]

Decision Explorer : <http://www.sagepub.co.uk/scolari/dec.exp.html>.

[Dieberger 1997]

Dieberger A. *Supporting social navigation on the World Wide Web*. International Journal of Human Computer Studies, 46 (1997), pp.805-825.

[Dieberger 1998]

Dieberger, A. *Social connotations of spatial metaphors and their influence on (direct) social navigation*, In Hook, K, Munro, A. and Benyon, D (Eds). Workshop on personalised and social navigation in information space, 16-17 March 1998. Swedish Institute of Computer Science, SICs Technical Report T98.2. Kista, Sweden (March 1998), pp.5 - 14.

[Grosser 1991]

Grosser, K. *Human networks in organizational information processing*. In M. Williams (Ed.) Annual Review of Information Science and Technology, 26, Medford, NJ: Learned Information Inc, 1991, 349 ff.

[Haythornthwaite 1996]

Haythornthwaite, C. *Social network analysis: an approach and technique for the study of information exchange*. LISR, 18 (1996), pp.323-342.

[Huotari 1998]

Huotari, M-L. *Social network analysis as a tool to evaluate IM in the public sector: a pilot study at the University of Tampere*. In Proceedings of ISIC 98, University of Sheffield, August 15-18, 1998. London: Taylor Graham (in press).

[Inspiration 1998]

Inspiration Software Inc: <http://www.teleport.com/~inspirat/coinfo.html>

[Lyons 1998]

Lyons, D. *The buzz about Firefly*. New York Times Magazine, June 29 1998, 37-40.

[Netmap 1998]

Netmap is a registered trademark of ALTA Analytics, Inc. <http://www.ALTAanalytic.com/>

[Rosenbaum 1997]

Rosenbaum, H. *Intranets and digital organizational information resources: towards a portable methodology for design and development*. In ASIS 97. Proceedings of the Annual meeting of the American Society for Information Science, 1997. Medford, NJ: Information Today Inc, 1997, 14-26.

[Scheepers and Damsgaard 1997]

Scheepers, R. and Damsgaard, J. *Using internet technology within the organisation: a structural analysis of intranets*. In Group 97. New York: ACM, (1997) pp.9-18

[Spence 1998]

Spence, R. *A framework for navigation. Information Engineering Section Report 98/2*. Department of Electrical and Electronic Engineering. Imperial College of Science, Technology, and Medicine: London, 1998.

[Sonnenwald 1998]

Sonnenwald, D. *Evolving perspectives of human behaviour: contexts, situations, social networks and information horizons*. In Proceedings of ISIC 98, University of Sheffield, August 15-18, 1998. London: Taylor Graham (in press).

[Tracey and Whittaker 1990]

Tracy, E.M. and Whittaker, *The social network map: assessing social support in clinical practice*. Families in Society, 71 (8), 1990, 461-470.

[Vishik 1997]

Vishik, C.M. *Internal information brokering and patterns of usage on corporate intranets*. In Group 97, Phoenix Arizona. New York: ACM, 1997, 111-118. (vishik@slb.com)

[Wexelblat and Maes 1998]

Wexelblat, A. and Maes, P. *Visualising histories for web browsing*. MIT Media lab, Software Agents Group, 1998. (wex.pattie@media.mit.edu)

"Telearbeit auf der Grundlage einer neuen Kommunikationstechnologie"

Patric Hoechst, Georg Schneider,
Karsten Merschjan, Jean Schweitzer

Patric Hoechst
Siemens Telekooperations Zentrum am DFKI¹
Stuhlsatzenhausweg 3
D-66123 Saarbrücken
E-Mail: Patric.Hoechst@stz.dfki.de

Inhalt

1. Einleitung und Begriffsklärung
2. Problemstellung und Vorgehensweise
3. Das Projekt „TeleArbeit Saar“
4. Sozialwissenschaftliche Begleitforschung
5. Zusammenfassung und Ausblick
6. Literatur

Zusammenfassung

Anhand des Projektes "TeleArbeit Saar" werden die Auswirkungen von Telearbeit auf Organisationen und Individuen der Informationsgesellschaft dargestellt. Das Papier beschreibt das Projektumfeld, die konkreten empirischen Fragestellungen sowie die Auswertung der einzelnen Faktoren.

1. Einleitung und Begriffsklärung

Der Übergang zur sogenannten Informationsgesellschaft wird einen ähnlich grundlegenden gesellschaftlichen Wandel auslösen, wie ihn die erste industrielle Revolution mit sich gebracht hat. Dies wird sich besonders deutlich auswirken, in bestehenden Arbeits- und Beschäftigungsstrukturen sowie im Verständnis der Menschen, was Arbeit überhaupt darstellt und wie sie organisiert wird [Van Haaren, Hensche 97].

¹ Das Siemens Telekooperations Zentrum (STZ) am Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) ist eine Forschungskooperation zwischen dem Zentralbereich Technik der Siemens AG und dem DFKI mit dem F&E-Schwerpunkt „Multimediale Telekooperation“. Mehr über das STZ siehe <http://www-stz.dfki.uni-sb.de>

Gerade Telearbeit bietet innovative Möglichkeiten und Lösungsansätze, Arbeit neu zu definieren und den veränderten Anforderungen der Informations- und High-Tech-Gesellschaft anzupassen.

Telearbeit wird als die wohnortnahe Arbeit unabhängig vom Firmenstandort an mindestens einem Arbeitstag pro Woche bezeichnet, wobei die (Zusammen-) Arbeit über räumliche Entfernungen hinweg unter primärer Nutzung von Informations- und Kommunikationstechnologien erfolgt und eine Telekommunikationsverbindung zum Arbeitgeber bzw. Auftraggeber zur Übertragung von Arbeitsergebnissen genutzt wird [Korte, Kordey 96]. Vereinfacht ausgedrückt, umfaßt der Begriff Telearbeit als Oberbegriff Tätigkeiten, die unterstützt durch Informations- und Kommunikationstechnik räumlich entfernt vom Standort des Arbeits- oder Auftraggebers durchgeführt werden [Godehardt 94].

Zwei verschiedene Ausprägungen der Telearbeit, mit denen wir uns hier beschäftigen, sind die "Teleheimarbeit" (TeleHome) und die "mobile Telearbeit" (TeleMobil).

Teleheimarbeit umschreibt ein Arbeitsverhältnis, bei dem der Arbeitnehmer von zu Hause aus arbeitet. Sie untergliedert sich zusätzlich nochmal in zwei unterschiedliche Bereiche: In „permanente“ und „alternierende“ Teleheimarbeit. Im Fall der permanenten Teleheimarbeit bedeutet das, daß der Telearbeiter keine festen Räumlichkeiten mehr in der Organisation hat. Er erhält zwar von der Firma seine Instruktionen, arbeitet jedoch durchgehend in seinem "Home-Office". Die "alternierende Teleheimarbeit" bezeichnet eher einen Kompromiß aus der neuen und der alten Arbeitsform. Hier arbeitet der Telearbeiter zwar auch grundsätzlich von zu Hause aus, hat aber immer noch ein bis drei Tage pro Woche, an denen er wie gehabt im Büro präsent ist, sowohl für Kunden als auch für Vorgesetzte und Kollegen. Seinen Büroraum teilt er in der Regel mit anderen Telearbeitern, und zwar nach einem genau festgelegten Plan, um Überschneidungen zu vermeiden. Bei der mobilen Telearbeit pendelt der Arbeiter zwischen betrieblicher Arbeitsstätte und dem Kunden mit seiner gesamten EDV-Anlage (in der Regel ein Laptop-PC), d.h. der Arbeitsplatz "wandert mit".

Telearbeit hat sowohl Befürworter als auch Gegner. Die Befürworter antizipieren einen enormen Anstieg an neuen Arbeitsplätzen [Rexrodt 97], also ein wirkungsvolles Mittel zur Bekämpfung des Problemthemas Arbeitslosigkeit, eine Stärkung der Wettbewerbsfähigkeit für den Standort Deutschland [Meyer 97]. Sie sehen mehr Freiräume und Selbstverantwortlichkeiten für Arbeitnehmer (Telearbeit fördert das Maturitätsbedürfnis), erwarten effektivere Arbeitsergebnisse, begrüßen, daß die Integration von erziehenden Elternteilen in das Arbeitsleben erhalten bleibt und vieles mehr.

Gegner befürchten jedoch eine zunehmende soziale Isolation der Arbeiter, geringere Arbeitseffizienz und mangelnde Kontrollmöglichkeiten. Aus diesem Grunde sollen die Auswirkungen von Telearbeit auf die betreffenden Arbeiter untersucht werden. Das Projekt "TeleArbeit Saar"² dient dabei als Umfeld für Überlegungen.

² Das Projekt "TeleArbeit Saar" wird von dem "saarländischen Ministerium für Wirtschaft und Finanzen" im Rahmen der Landesinitiative "Telekommunikation Saar" gefördert. Projektpartner sind die "Beratungsstelle für sozialverträgliche Technologiegestaltung" (Best e.V.) der

2. Problemstellung und Vorgehensweise

Das Untersuchungsinteresse geht prinzipiell dahin, daß zunächst einmal als erste Problemstellung der Produktivitätsgewinn der neuen Arbeitsform zu betrachten ist. Hier wird die generelle Akzeptanz von Telearbeit als auch darauf aufbauend die Akzeptanz der Kommunikationstechnologie erhoben. Weiterhin werden Veränderungen in den Arbeitsgewohnheiten, den Arbeitsprozessen, der Kooperation und persönliche Faktoren wie Arbeitszufriedenheit und Motivation betrachtet und ausgewertet. Dieser Teil hebt sich von vorangegangenen Untersuchungen [Glaser 95] insofern ab, als daß nicht nur der Telearbeiter per se, sondern alle Hierarchieebenen des Management, die betrieblichen Interessenvertretungen und unterschiedliche organisationelle Interessenschwerpunkte in die Untersuchung eingebunden werden. Das spezielle Augenmerk richtet sich hier auf den Kontrollverlust, den viele Manager der mittleren Verwaltungsebenen propagieren, wenn es um das Thema Telearbeit geht.

Über Voraussetzungen und Kompetenzen von Arbeitern, die an dieser neuen Form von Arbeit partizipieren wollen, ist in älteren Untersuchungen nie eine klare Aussage getroffen worden. Somit bestehen bis Heute noch Probleme bei der Personalauswahl. Als eine weitere Problemstellung wird nun zusätzlich die Profilerstellung des "geeigneten Telearbeiters" integriert.

Der dritte Schwerpunkt beinhaltet die Identifizierung des Arbeiters mit der Organisation. Ein bekanntes Problem war bisher immer, daß durch Outsourcing-Prozesse, wenn also Organisationen ganze Strukturen nach außen verlagert haben, für die Betroffenen Probleme mit ihren Identifizierungsmöglichkeiten in Richtung Organisation entstanden. Hier gilt es einen Nachweis zu erbringen, daß der Einsatz modernster Kommunikationstechnologie und das damit gezeigte Vertrauen in die Kompetenzen des Mitarbeiters eine neue Form der "Corporate Identity" im Bereich Telearbeit ermöglicht.

Letztlich soll nach Beendigung des Projektes ein Erhebungsinstrument, welches unabhängig von einzelnen Organisationszielen- und -ausprägungen im Bereich Telearbeit eingesetzt werden kann, für alle interessierten Forscher zur Verfügung stehen.

3. Das Projekt "TeleArbeit Saar"

Umfeld

Innerhalb des Projektes "TeleArbeit Saar" wurde ein innovativer und ganzheitlicher Ansatz gewählt [Merschjan 97]. Dieser bündelt ebenso die Kompetenzen auf dem Gebiet der für die Telearbeit notwendigen Einführungskonzepte, -methoden und der dafür benötigten technischen Systeme, als auch das unerläßliche Know-how für eine erfolgreiche arbeitswissenschaftliche Begleitforschung. Als herausragende Items sind zu nennen:

Arbeitskammer des Saarlandes, der Versicherungsanbieter "CosmosDirekt", das Beratungs- und Softwarehaus "IDS Prof. Scheer", das "Ministerium für Frauen, Arbeit, Gesundheit & Soziales" sowie das "Siemens Telekooperations Zentrum" (STZ).

- ein generelles Einführungsmodell für Telearbeit
- ein arbeitswissenschaftlicher Leitfaden
- ein allgemein verfügbares Erhebungsinstrument
- der integrierte Telearbeitsplatz
- das Tele-Stress Modell [Brandstätter 98]
- die Gründung des TACC (Saarländisches TeleArbeits-Competence-Center)

Kommunikations-Technologie

Die im Projekt verwendete Kommunikationstechnologie ist das "I-View" der Siemens AG, ein Desktop-Videoferenzsystem, welches auditive und visuelle Informationen von den einzelnen Gesprächs- oder Konferenzpartnern liefert. Die Benutzerschnittstelle wird durch "Group-Win" realisiert, eine STZ-eigene Softwarekomponente, die auf I-View aufsetzt. Damit ist ein komfortabler Zugang zu den Audio- Videokanälen einschließlich Application-Sharing möglich. Die einzelnen Kommunikationspartner können gemeinsam und "online" (synchron) an ein und demselben Dokument arbeiten. Diese Technologie ermöglicht dem Telearbeiter, neben dem bisher bei dieser Arbeitsform verfügbaren auditiven nun auch den weitaus bedeutsameren visuellen Informationskanal zu integrieren und wenn erforderlich Arbeiten zeitgleich im TeleTeam zu erledigen.

Telearbeits-Szenarien

Telearbeit wird bei den verschiedenen Projektpartnern auf unterschiedlichste Weise durchgeführt - immer an die individuellen Bedürfnisse der Organisation angepaßt.

- *Ortsunabhängige Policing* – Dieses Anwendungsszenario wird vom Versicherungsanbieter "CosmosDirekt" umgesetzt. Die Telearbeiter führen von zu Hause aus (TeleHome) die Policingen durch und verschicken die fertig bearbeiteten Kontrakte per Kommunikationsmedium zum Arbeitgeber. Weiterhin haben Sie die Möglichkeit, auf die umfangreiche Datenbank des Versicherungskonzerns per ISDN zuzugreifen und stehen auch gleichzeitig mit dem Rückversicherer in Verbindung, sollten Probleme auftauchen. In diesem Szenario wird die alternierende Telearbeit umgesetzt.
- *Teleconsulting* - Das Software- und Consultinghaus "IDS Prof. Scheer" arbeitet auf der Grundlage der mobilen Telearbeit (TeleMobil). Hier sind die Arbeiter nicht von zu Hause aus tätig, sondern pendeln zwischen betrieblicher Arbeitsstätte und Kunden. Beim Kunden vor Ort wird die neue Kommunikationstechnologie eingesetzt um Hilfestellungen und Informationen bei konkreten Problemstellungen im Bereich Technik oder Consulting einzuholen. Zuvor wurde dieser Vorgang über umständliche Wege umgesetzt, wobei es hin und wieder zu mehreren Terminen kam. Dies wird jetzt mittels mobiler Telearbeit in einem einzigen Zuge verwirklicht.
- *Interministerielle Telekooperation* - Dieses dritte Szenario unterstützt die Arbeiten auf ministerieller Ebene (Ministerium für Frauen, Arbeit, Gesundheit und Soziales) und der Staatskanzlei der saarländischen Landesregierung.

Beide Strukturen sind per ISDN miteinander vernetzt, wobei die einzelnen Referenten zwischen dem Amt der Ministerin und der Staatskanzlei schnell und unproblematisch Aufgaben in Minuten gemeinsam bearbeiten und durchsprechen können, welche durch die notwendigen Postwege ansonsten Tage und Wochen benötigen würden. Die Telearbeiter bleiben hier wie gehabt an ihrem normalen Arbeitsplatz, werden aber durch die Technologie (Vidokonferenz und Application-Sharing) in ihrem Arbeitsprozeß unterstützt.

4. Sozialwissenschaftliche Begleitforschung

Der sozialwissenschaftlichen Begleitforschung, die in Zusammenarbeit mit der Best e.V. durchgeführt wird, liegt die Erstellung eines sozialen Leitfadens für die Projektdurchführung (Bewahrung bestehender Beschäftigungsverhältnisse, Klärung arbeitsrechtlicher Fragen, Einbindung der Interessenvertretungen etc.) und eines ad hoc konstruierten Erhebungsinstrumentes zugrunde, welches auch für zukünftige Untersuchungen im Bereich Telearbeit zur Verfügung stehen wird.

Erhebungsfaktoren

Inhaltlich werden hierbei folgende Faktoren erhoben:

1. *Statistische Werte* - Anhand von Geschlecht, Alter und anderen persönlichen Daten soll ein Überblick gewonnen werden, ob geschlechtliche oder altersbedingte Faktoren bei der Durchführung von Telearbeit eine Rolle spielen.
2. *Arbeitsbeschreibung, Einschätzung der eigenen Arbeit und Arbeitsprozesse* - Mit der Erhebung dieser Faktoren sollen Tätigkeiten auf ihre Tauglichkeit für Telearbeit hin untersucht und eventuell neue Bereiche sondiert werden, die für herkömmliche Telearbeit bislang ungeeignet waren.
3. *Formelle und informelle Kommunikation* - Diese Faktoren dienen dazu, erst einmal zu erheben, wie gut die einzelnen kommunikativen Strukturen gestaltet sind, ob sie sich für die Durchführung von Telearbeit eignen und letztendlich wie effektiv sie genutzt werden können, um die Kommunikation mit der Organisation zu gewährleisten. Weiterhin wird betrachtet, ob die Telearbeiter zu stark isoliert werden und somit der Vorwurf der sozialen Isolation gerechtfertigt ist.
4. *Introversion/Extraversion* - Hier soll eindeutig die alte Hypothese widerlegt werden, nach der introvertierte Personen über die beste Eignung zum Telearbeiter verfügen.
5. *Organisationsklima* - Bei diesem Faktor wird untersucht, ob Telearbeit positive oder negative Auswirkungen auf das Arbeitsklima mit sich bringt und ob im ersten Fall die Telearbeit dazu dienen kann, ein schlechtes Klima zu verbessern.
6. *Arbeitszufriedenheit* - Hier werden ähnlich wie beim letzten Punkt die Auswirkungen von Telearbeit auf essentielle Fragestellungen der Organisationspsychologie erhoben.
7. *Arbeitsmotivation* - Mit diesem Faktor werden die Auswirkungen der neuen Arbeitsform und Kommunikationstechnologie auf die Arbeitsmotivation der Telearbeiter evaluiert.
8. *Erwartungen, Befürchtungen und Erfahrungen mit Telearbeit* - Die Telearbeiter werden ganz spezifisch über ihre Erwartungen und Befürchtungen

mit dieser neuen Arbeitsform befragt und in späteren Erhebungen ihre bestehenden Erfahrungswerte erhoben und analysiert.

9. *Technische Evaluation* - Die neue Kommunikationstechnologie wird bezüglich ihrer Qualität, Stabilität und Akzeptanz hin überprüft, wobei die einzelnen Komponenten wie Video, Audio und Application-Sharing separat betrachtet werden.

Beispiel

Exemplarisch soll hier noch einmal auf den vorangegangenen Punkt 4) eingegangen werden.

Versucht wird anhand dieser Skalen die von Glaser in seiner Studie [Glaser 95] schon angedeutete Tendenz, daß eben *nicht* wie bisher immer vermutet der introvertierte Mensch als geeigneter Telearbeiter in Frage kommt, da er dem Verlust kommunikativer Strukturen eher gleichgültig gegenübersteht, zu betrachten. Wir vertreten den Standpunkt, daß ein extravertierter Mensch sich besser zum Telearbeiter eignet, da er von sich heraus in der Lage ist, die Kommunikation zu den Kollegen und der Organisation aufzubauen und aufrechtzuerhalten, somit erst gar nicht die Problematik einer sozialen Isolation aufkommt.

Die im Rahmen des Projektes eingesetzte Kommunikationstechnologie soll diesen Prozeß zusätzlich noch verstärken. Gerade durch die unproblematische und schnelle Möglichkeit eines kommunikativen Verbindungsaufbaus per Videokonferenz mit den Kollegen in der Organisation werden die extravertierten Telearbeiter in ihren Bemühungen noch unterstützt.

Andererseits kann der einfache Zugang zum Kommunikationsmedium die "Nachteile" des introvertierten Telearbeiters egalisieren, die Glaser-These [Glaser 95] also falsifizieren. Dies wird im weiteren Verlauf der Evaluation genauer geprüft werden. Dies geschieht mit Fragen, die bei den Probanden eine Zuordnung in die Persönlichkeits-Klassifikationen "Introversion" oder "Extraversion" ermöglichen. Die Ausprägungen der Telearbeiter auf dieser Dimension werden erhoben, wobei die gewonnenen Werte anschließend mit den Ausprägungen der einzelnen Versuchspersonen in den Faktoren "Arbeitszufriedenheit" und "Motivation" verglichen werden. Der Unterschied zu vorangegangenen Untersuchungen liegt wie oben bereits erwähnt in der Tatsache, daß alle Hierarchieebenen dazu betrachtet werden. Die I/E-Fragen können folgendermaßen konzipiert sein:

Ich habe gerne mit Aufgaben zu tun, die schnelles Handeln erfordern.

stimmt nicht

stimmt

1	2	3	4	5
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Erhebungsrahmen

Die Evaluation findet in vier Schritten statt. Nach der Erhebung der aktuellen "Ist-Situation" in den jeweiligen Organisationen (sogenannte Nullmessung, November 97 bis Januar 98) werden drei aufeinanderfolgende Feldmessungen durchgeführt (April/Mai 98, Juni 98, Juli/August 98). Der Abstand zwischen den

einzelnen Messungen wird mindestens einen Monat betragen, um eventuelle Veränderungen in persönlichen oder organisationellen Faktoren entdecken zu können. Erhoben werden eine Treatment-Gruppe (Telearbeiter) und eine Kontrollgruppe (Mitarbeiter, die weiterhin in ihrer herkömmlichen Arbeitsform fortfahren), um einen eindeutigen Bezugsrahmen zu schaffen. Darauf aufbauend werden dann zum einen die Entwicklungstendenzen der Telearbeiter über die einzelnen Meßzeitpunkte hinweg (Varianz innerhalb der Treatment-Gruppe) und zum anderen die Unterschiede zwischen Treatment- und Kontrollgruppe zu den einzelnen Meßzeitpunkten (Varianz zwischen den Gruppen) erhoben. Die Varianz innerhalb der Kontrollgruppe wird noch zusätzlich erfaßt, um eventuelle Schwankungen in der herkömmlichen Arbeitsform zu registrieren.

Das Erhebungsinstrument besteht aus 233 einzelnen Fragen, die nach den oben erwähnten 9 Faktoren einzuordnen sind. Der größte Teil dieser Fragen ist explizit für dieses Erhebungsinstrument erstellt worden, ein weiterer Teil ist aus [Glaser 95] entnommen. Dieser Schritt war notwendig, da verschiedene wirtschaftliche Kompetenzen und Betätigungsfelder in *einem* Projekt integriert wurden und diesbezüglich noch kein geeignetes und allgemein zugängliches Instrument zur Verfügung stand.

Bei den Fragen werden drei verschiedene Ausprägungen unterschieden. Zum einen die bereits im Beispiel Introversion/Extraversion gezeigten Fragen auf der Grundlage der "Likert-Skala" (siehe oben) mit Polarisierungen von 5 oder 7 Punkten, unskalierte Entscheidungsfragen mit mehreren Antwortalternativen

Besteht die Möglichkeit, bei der Durchführung ihrer Arbeitspakete auf die Hilfe ihrer Kollegen zurückzugreifen?

- ☐ Ja
- ☐ Nein

oder ganz offene Fragen.

Wieviele Stunden arbeiten Sie tatsächlich in einer Woche?

5. Zusammenfassung und Ausblick

Zum Stand des Projektes läßt sich sagen, daß bis dato die Nullmessung und die erste von drei Feldmessungen abgeschlossen sind. Hier zeigt sich bereits die Tendenz, daß zum einen die Bedeutung informeller Kommunikationsstrukturen sehr hoch bewertet wird, wenngleich nicht von einem Mangel auszugehen ist, und zum anderen ein gewisser Erfahrungseffekt mit der Zeit eintritt, daß also zum einen zu hohe Erwartungen zurückgeschraubt werden, zum anderen aber auch simultan Befürchtungen zurückgehen. Ausführlichere Befunde können jedoch erst nach Abschluß der letzten Feldmessung gewonnen werden.

In einer weitergehenden Untersuchung sollen Belege dafür gefunden werden, daß Telearbeit viele Problemfaktoren herkömmlicher Arbeitsstrukturen aufheben kann und die neue Kommunikationstechnologie zusätzliche Arbeitsprozesse in das Spektrum "Telearbeit" integrieren wird, da sie die her-

kömmlichen Telearbeits-Kommunikationswege wie Telefon oder E-Mail durch visuelle Informationen ergänzt - also mehr Persönlichkeit und Qualität in die Interaktion bringt, da mehr Informationskanäle zur Verfügung stehen, die vom Telearbeiter zur Informationsverarbeitung herangezogen werden können (siehe hierzu auch[Reichwald 98]).

Abschließend möchten wir uns bei Jörg Steinhausen (STZ) und Dr. Bernd Grass (Best e.V.) für die Unterstützung und Beratung bedanken.

6. Literatur

[Van Haaren, Hensche 97]

Van Haaren, K., Hensche, D., „Arbeit im Multimedia-Zeitalter“, VSA-Verlag Hamburg, (1997)

[Korte, Kordey 96]

Korte, W., Kordey, N., „Telearbeit erfolgreich realisieren“, Vieweg., (1996)

[Godehardt 94]

Godehardt, B., "Entwicklung der Telearbeit und Begriffsbestimmungen", BMWi, (1997)

[Rexrodt 97]

Rexrodt, G., "Expertenmeinungen zum Thema Telearbeit", in "teleworx - Das Magazin für Telearbeit und Telelearning", Ausgabe Januar 1997, IWT-Verlag, (1997)

[Meyer 97]

Meyer, U., "Expertenmeinungen zum Thema Telearbeit", in "teleworx - Das Magazin für Telearbeit und Telelearning", Ausgabe Januar 1997, IWT-Verlag, (1997)

[Glaser 95]

Glaser, W., „Telearbeit in der Praxis - Psychologische Erfahrungen mit Außerbetrieblichen Arbeitstätten bei der IBM Deutschland GmbH“, Neuwied, (1995)

[Merschjan 97]

Merschjan, K., TeleArbeit Saar, http://www-stz.dfki.uni-sb.de/stz/projects/telearbeit_saar/index.html, (1997)

[Brandstätter 98]

Brandstätter, S., "Tele-Stress: Motor oder Bremse der (persönlichen) Produktivität", Diplomarbeit am Siemens Telekooperations Zentrum, (1998)

[Reichwald 98]

Reichwald, R. et al., „Telekooperation, Verteilte Arbeits- und Organisationsformen“, Springer Verlag, (1998)

Wissensmanagement in Lernenden Organisationen

Eine neue Herausforderung für die Informationswissenschaft?

Otto Krickl, Elisabeth Milchrahm

Prof. Dr. Otto Krickl
Institut für Betriebswirtschaftslehre
der öffentlichen Verwaltung und
Verwaltungswirtschaft
Karl-Franzens-Universität Graz

Universitätsstraße 15, ReSoWi E4
A-8010 Graz
email: otto.krickl@kfunigraz.ac.at

Univ.-Ass. Mag.
Elisabeth Milchrahm
Institut für Informationswissenschaft
Karl-Franzens-Universität Graz

Universitätsstraße 15, ReSoWi F3
A-8010 Graz
email:
elisabeth.milchrahm@kfunigraz.ac.at

1. Inhalt

1.	Inhalt	1
2.	Zusammenfassung	2
3.	Einleitung	2
4.	Zielsetzung	3
5.	Lernenden Organisationen	3
5.1.	Charakteristika	4
5.2.	Bezugsrahmen zur Einordnung der informationswissenschaftlichen Aufgaben	4
6.	Groupware-unterstützte Projektarbeit	7
6.1.	Typische Formen der Projektabwicklung	7
6.2.	Methoden des Projektmanagements	8
6.3.	Groupwareunterstützung bei der Projektabwicklung	9
7.	Informationsquellenanalyse	9
7.1.	Organisationale Routinen	9
7.2.	Organisationale Handlungstheorie	10
7.3.	Informationssystem	11
7.4.	Störungs-Reaktions-Verknüpfungen	11
8.	Informationsaufbereitung	11
9.	Schlußbetrachtung	13
10.	Literaturverzeichnis	13

2. Zusammenfassung

In der sehr umfangreichen Literatur zum Konzept der Lernenden Organisation werden die vielfältigen Potentiale zur Effizienzsteigerung und deren Bedeutung für die Wettbewerbsfähigkeit herausgestellt. Basis dafür bildet unter anderem der systematische Aufbau einer Wissensdatenbank und deren Nutzung in Verbindung mit den verschiedenen Instrumenten der Lernenden Organisation. Für die Darstellung des Aufbaues von Wissensdatenbanken findet jedoch der aktuelle Stand der Forschung im Bereich der Informationswissenschaft keine ausreichende Berücksichtigung. Da aus diesem Forschungsgebiet aber wesentliche Impulse für die Realisierung des Konzeptes der Lernenden Organisation gewonnen werden können, soll im vorliegenden Beitrag die Basis für ein Methodenset für informationswissenschaftliche Aufgaben in diesem Kontext vorgestellt werden. Dabei wird aufgezeigt, daß die Informationsaufbereitung eine fallspezifisch konzipierte Methodik voraussetzt. Details der Konzeption des Methodensets werden im Vortrag im Rahmen der Tagung vorgestellt. Eine Evaluierung der Ergebnisse in der Praxis ist vorgesehen.

3. Einleitung

Lernende Organisationen sind in der Praxis noch sehr selten und zumeist nur ansatzweise realisiert. Die Gründe für die relativ geringe Umsetzung sind vor allem in der großen Unbestimmtheit der Konzepte zu finden, da konkrete Realisierungsfragen¹ in der Literatur nur selten behandelt werden. Zumeist sind es Unternehmensberater², die einzelne Realisierungen darstellen, diese aber auch als Dienstleistung anbieten und daher in der allgemein zugänglichen Literatur nicht konkret beschreiben.

Für den Gestalter in der Praxis ergibt sich daher die unbefriedigende Situation, daß zwar die großen Vorteile der Lernenden Organisation hervorgehoben werden, jedoch kaum Material zur konkreten Umsetzung zu Verfügung steht.

Dieses Informationsbedürfnis wird auch von der Informationswissenschaft nicht voll abgedeckt. Bei der Beschreibung der grundlegenden Methoden finden sich beispielsweise Hinweise auf die Notwendigkeit einer Handlungsrelevanz und des Einbindens von Informationen in einen organisatorischen Kontext.³ Die dahinter stehende Methodik wird jedoch nicht im Detail beschrieben und daher nicht intersubjektiv nachvollziehbar. Andererseits zeigen Fallbeispiele das eine

¹ Hier sind die Informationsquellenanalyse, die Methoden der Informationsaufbereitung und die laufende Anpassung der Wissensbasis an Umweltveränderungen gemeint.

² Vgl. beispielsweise Arthur D. Little International: siehe dazu Ortwein und Spallek 1998, S. 105 ff.

³ Vgl. beispielsweise Kuhlen, der diesen Aspekt als Kriterium für die Akzeptanz für die Vermittlung von Informationsdienstleistungen definiert: Kuhlen 1996, S. 355.

systematische Informationsaufbereitung mit fallspezifisch entwickelten Methoden (Penker und Proske 1993, S. 687 f.) optimale Ergebnisse erbringen.

4. Zielsetzung

Zielsetzung der Arbeit ist einen Beitrag zum verstärkten Forschungsbedarf im Bereich der Identifizierung von konkreten Maßnahmen und Methoden zu leisten, die ein effizientes Wissensmanagement in Lernenden Organisationen sicherstellen. Als Szenario für die Darstellungen wird die groupware-unterstützte Projektabwicklung herangezogen, da hier die größten Potentiale für den Wissenstransfer innerhalb und zwischen Projekten bestehen. Die Generierung von neuem expliziten Wissen aus Projektdokumentationen und der systematische Aufbau eines Kontextbezuges als Basis von Wissensnetzwerken⁴ bilden daher den Schwerpunkt der Arbeit.

Die Basis dafür bildet ein Bezugsrahmen, der die Eingliederung der informationswissenschaftlichen Aufgaben in das Gesamtkonzept der Lernenden Organisation beinhaltet. Darauf aufbauend wird ein Szenario für projektorientierte Aufgabenabwicklung dargestellt, die durch Groupwaretools unterstützt wird. Die dabei typischerweise anfallenden Informationen werden analysiert und auf ihre Eignung zur Entwicklung von Methoden zur Aufbereitung des Wissens für die Wissensdatenbank untersucht.

5. Lernenden Organisationen

In der westlichen Managementtradition wird Wissen als etwas Formales, Systematisches und damit in Worten und Zahlen Darstellbares aufgefaßt. Wissen wird damit explizit, da es problemlos von einem zum anderen Menschen weitergegeben und mittels IT verarbeitet werden kann. Japanische Unternehmen haben demgegenüber ein völlig anderes Verständnis von Wissen, da hier Wissen hauptsächlich als etwas implizites angesehen wird. Dieses ist persönlich, entzieht sich dem formalen Ausdruck und läßt sich nur schwer mitteilen (Nonaka und Takeuchi 1997, S. 18 f.). Implizites Wissen wird hier auch in zwei Dimensionen unterteilt. Einerseits die technische Dimension, die Fertigkeiten (manueller und informeller Art) umfaßt. Andererseits wird die kognitive Dimension, die mentale Modelle und Vorstellungen umfaßt, unterschieden.

Aus der Erkenntnis des impliziten Wissens wurde in japanischen Unternehmen das Konzept der Lernenden Organisation entwickelt. Dieses geht von dem Grundprinzip aus, daß Wissen nur von Einzelpersonen geschaffen werden kann, da eine Organisation ohne den einzelnen kein Wissen schaffen kann. Die organisatorischen Strukturen sind jedoch so zu gestalten, daß kreative

⁴ Zum Aufbau von Wissensnetzwerken vergleiche beispielsweise: Servatius 1998, S. 103.

Personen bei der Wissensschaffung unterstützt werden. Da aber implizites Wissen für die Organisation ohne weitere Gestaltungsmaßnahmen nicht nutzbar ist, werden im nächsten Abschnitt die für die Problemstellung relevanten Aspekte der Lernenden Organisation im Überblick dargestellt.

5.1 Charakteristika

Ein weithin akzeptiertes Konzept bzw. eine entsprechende einheitliche Begriffsvorstellung Organisationalen Lernens existieren bisweilen nicht. Aus den zahlreichen Definitionen und Ansätzen (vgl. beispielsweise Senge 1990) zur Theorie der Lernenden Organisation läßt sich folgende Arbeitsdefinition für den Begriff des Organisationalen Lernens skizzieren (vgl. insbesondere Probst und Büchel 1998, S.17):

Unter Organisationalem Lernen ist der immerwährende Prozeß der Veränderung der Wert- und Wissensbasis auf organisationaler Ebene, der im Wechselspiel zwischen Individuen und Organisation abläuft und der bezogen auf die interne und/oder externe Umwelt der Organisation stattfindet, die Verbesserung der Problemlösungs- und Handlungskompetenz, sowie die Veränderung des gemeinsamen Bezugsrahmens von und für Mitglieder innerhalb einer Organisation zu verstehen.

Lernende Organisationen zeichnen sich dadurch aus, daß sie Informationen und Wissen nicht nur effizient verarbeiten, sondern darüberhinaus eigenständig hervorbringen.

Grundvoraussetzung für den fortwährenden Prozeß der organisationsweiten Wissensschaffung ist die Nutzung des impliziten Wissensbestandes der einzelnen Organisationsmitglieder. Entscheidende Bedeutung kommt dabei dem Aufbau und der Pflege einer Wissensdatenbank zu. Die Wissensdatenbank als kollektives Gedächtnis der Organisation generiert relevantes organisationales Wissen und stellt dieses Wissen durch Vernetzung den Mitarbeitern zur Verfügung.

5.2 Bezugsrahmen zur Einordnung der informationswissenschaftlichen Aufgaben

Voraussetzung für ein ganzheitliches Wissensmanagement ist, daß die Organisation über die Fähigkeit verfügt, die angebotenen Informationsmengen durch Anreicherung von Wissen zur Problemlösung zu nutzen.

Die zweckorientierte Vernetzung von Informationen gekoppelt mit kontextabhängigen Wissen ermöglicht die Entwicklung neuen Wissens und trägt in Folge zum Aufbau einer organisationalen Wissensbasis bei.

Um Wissen kommunizierbar zu machen und organisationales Wissen zu generieren, postulieren Nonaka und Takeuchi (1997, S. 74-84) vier Formen der Wissensumwandlung:

Sozialisation, der Austausch von implizitem Wissen, beruht auf der Annahme, daß durch Beobachtung von Kollegen und Nachahmung des Beobachteten organisationsbezogenes implizites Wissen übertragen wird. Der Blickwinkel erfolgt hier aus der Sicht des Wissenstransfers in Verbindung mit dem zugehörigen Erfahrungskontext.

Externalisation, die Umwandlung von implizitem in explizites Wissen, stützt sich auf den Gebrauch von Metaphern, Analogien und Modellen.

Kombination dient der Verbindung verschiedener Bereiche von explizitem Wissen zur Generierung neuen Wissens.

Internalisierung bezeichnet den Prozeß der Eingliederung expliziten Wissens in das implizite Wissen. Nonaka schlägt vor, Wissen in Dokumenten festzuhalten, um das Sammeln von Erfahrungen und die Übermittlung expliziten Wissens zu erleichtern.

Diese vier Formen der Wissensgenerierung bzw. -umwandlung stehen in einer dynamischen Beziehung zueinander und fügen sich im Idealfall zu einem selbstverstärkenden Prozeß zusammen.

Der Prozeß zur unternehmensinternen Wissensgenerierung wird im Überblick durch ein Phasenmodell deutlich (siehe Abb.1). Dabei ist auf eine ganzheitliche Betrachtungsweise unter Berücksichtigung der sich ergebenden Wechselbeziehungen zwischen den einzelnen Phasen zu achten.

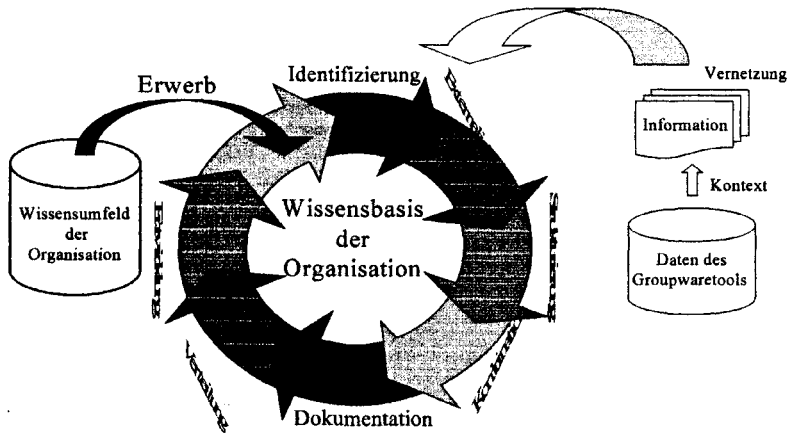


Abb.1: Phasenmodell des Wissenmanagements
(modifiziert nach Rey u.a. 1998, S. 31)

Der Kreislauf wissensorientierter Prozesse beginnt typischerweise mit dem Erwerb unternehmensexternen Wissens. Im Zuge der Wissensidentifizierung werden unternehmensexterne und -interne Wissensquellen erkundet (z.B. durch Expertenverzeichnisse, Wissenskarten), um als Resultat relevantes Wissen dem Einzelnen in der Organisation transparent zu machen.

Das identifizierte implizite Wissen wird im Rahmen der Externalisierung durch Verwendung von Metaphern, Analogien und Modellen in explizites Wissen umgewandelt. Dabei decken Metaphern Widersprüche impliziten Wissens auf, die durch Analogien zu harmonisieren und abschließend in Modellwissen umzuformulieren sind. (Nonaka und Takeuchi 1997, S. 80).

In der Phase der Wissensstrukturierung erfolgt dann die Bewertung des Wissens hinsichtlich Relevanz, Aktualität, Verfügbarkeit etc. Die Bewertungskriterien orientieren sich dabei an den Bedürfnissen der Informationsnutzer.

In der Kombinationsphase wird das strukturierte explizite Wissen verschiedener Unternehmensbereiche miteinander verbunden. Die Kombination vorhandenen expliziten Wissens kann so zu neuem Wissen führen.

Im Rahmen des Dokumentationsprozesses werden Erschließungs- und Suchstrategien zur Nutzbarmachung des extrahierten Wissens aufgestellt.

Das dokumentierte Wissen wird in Folge über Wissensnetzwerke durch Kommunikations- und Logistikprozesse auf die Organisationsmitglieder verteilt.

Dabei ist besonders auf eine anforderungs- und anwendungsgerechte Aufbereitung des Wissens zu achten.

In der letzten Phase der Wissensentwicklung wird über einen erneuten Durchlauf des Phasenmodells das dokumentierte Wissen aktualisiert und erweitert. Im Zusammenhang mit definierten Wissenszielen (vgl. Probst u.a. 1998, S. 63-95), die den Erfolg bzw. Mißerfolg von Wissensmanagement überprüfbar machen sollen, ist es Aufgabe der Wissensentwicklung nach Abschluß eines Projektes die Prozeßergebnisse kritisch zu hinterfragen.

Anhand des vorgestellten Phasenmodells wird die Bedeutung einer permanenten Weiterentwicklung organisationalen Wissens durch Selektion und Speicherung organisationsrelevanter Ereignisse bzw. Prozesse aufgezeigt. Der Kreislaufcharakter des Modells unterstützt die Integration von Wissen in die organisationale Wissensbasis und trägt so zur schrittweisen Institutionalisierung des Wissensmanagements bei.

6. Groupware-unterstützte Projektarbeit

Groupware-Systeme⁵ unterstützen die Kooperation von üblicherweise interdisziplinär zusammengesetzten Projektgruppen, die Aufgaben mit schwachen Strukturierungsgraden und geringer Vorhersehbarkeit zu lösen haben. Derartige wissensintensive Projektarbeiten haben meist Einmaligkeitscharakter (Aufgabenstellung ist zumindest teilweise neu) und erfordern einen koordinierten Wissenstransfer, da das zur Problemlösung vorhandene Fachwissen auf die Mitwirkenden verteilt ist.

Um das der Arbeit zugrunde liegende Verständnis der Anteile von impliziten bzw. expliziten Wissen im Datenbestand eines Groupwaretools transparent zu machen, werden in den nächsten Abschnitten einige Aspekte der Projektarbeit dargestellt.

6.1 Typische Formen der Projektabwicklung

Beim Start von Projekten liegt zumeist weder eine klare Analyse der Istsituation noch ein operationalisierter Zielkatalog vor. Um die Projektarbeit zumindest grob zu strukturieren wird in vielen Unternehmen ein Vorgehensmodell zur Projektabwicklung als verbindlich erklärt. Vorgehensmodelle untergliedern die Projektabwicklung in einzelne Phasen, wobei für jede Phase die hier zu leistenden Arbeitsschritte definiert sind. Ein wichtiges Element von Vorgehensmodellen sind auch die Abstimmungszeitpunkte. D. h. es wird festgelegt welche Informationen das Projektteam zu welchem Zeitpunkt der Projektabwicklung vorzulegen hat. Klargestellt wird dabei auch wie das

⁵ Es existiert keine allgemein akzeptierte Definition dessen, was man unter Groupware zu verstehen hat. Zur Kategorisierung von Groupware-Systemen vergleiche zum Beispiel: Österle 1996 S. 55 ff., Wagner 1995 S. 73 ff.

Entscheidungsgremium zusammengesetzt ist und worüber in den einzelnen Phasen entschieden wird.

Die Richtlinien für die Abwicklung von Projekten sind häufig nur teilweise als explizites Wissen verfügbar. Dies begründet sich auch damit, daß gerade in der Projektarbeit die Weitergabe durch Sozialisation erfolgt. Beispielsweise ist es oft üblich, vor der offiziellen Sitzung eines Entscheidungsgremiums mit bestimmten Personen die anstehenden Entscheidungen abzustimmen. Diese organisatorischen Routinen sind in den seltensten Fällen explizit dokumentiert, obwohl diesen eine starke Handlungsleitung zukommt.

Da Projekte zumeist nur mit kurz gefaßten Protokollen (bzw. Berichten, Lösungsbeschreibungen etc.) dokumentiert werden, kann aus dem expliziten Wissen nur ein Bruchteil des tatsächlichen Wissens entnommen werden. Insbesondere die sozialen Prozesse, die zu bestimmten Entscheidungen geführt haben, unterliegen kaum einer Externalisation. Da aber oft Entscheidungen nur vor dem Kontext des sozialen Prozesses während des Zustandekommens verständlich sind, fehlen die kognitiven Operatoren⁶ zur Einordnung der verfügbaren expliziten Wissensbestandteile.

6.2 Methoden des Projektmanagements

Methoden des Projektmanagements werden zur Planung und Steuerung der Projektarbeit angewendet. Beispielsweise werden Projektstrukturpläne als Basis für die Aufgabenverteilung und die Zeit-, Kosten- und Terminplanung verwendet.

Von besonderem Interesse sind aber im Zusammenhang mit der Problemstellung, die Methoden zur Erhebung von Ist-Zuständen, zur Entwicklung von Lösungsvorschlägen oder zur Entscheidungsfindung. Problemstrukturierungs-, Kreativitäts- und Bewertungstechniken werden in unterschiedlichen Varianten angewendet, sollten aber aus der Dokumentation auch Rückschlüsse auf den Entwicklungsprozeß zulassen – Präzisierung dazu siehe im nächsten Abschnitt.

⁶ Zum Begriff der kognitiven Operatoren siehe Kapitel 8.

6.3 Groupwareunterstützung bei der Projektabwicklung

Die Groupwareunterstützung bei Projektabwicklung umfaßt die gesamte Kommunikation und Erstellung bzw. Bearbeitung von Dokumenten und Daten. Darüber hinaus bieten gute Groupwaretools auch Funktionen zur Anwendung von Kreativitätstechniken (z.B. Brainstorming), Clusterbildung oder Abstimmungsprozessen.

Im Gegensatz zur konventionellen Projektabwicklung lassen sich aus den Daten des Groupwaretools wesentlich mehr Auswertungen erstellen, da auch der Entwicklungsprozeß dokumentiert wird. D. h. die gesamte Projektkommunikation, die Versionen verschiedener Dokumente, die Autoren einzelner Textpassagen etc. sind verfügbar und erlauben somit einen wesentlich tieferen Einblick in den Projektablauf.

Dennoch kann die Datenbasis des Groupwaretools nicht als Wissensdatenbank angesehen werden, da die einzelnen kognitiven Elemente⁷ nicht aufbereitet und nicht in Beziehung zueinandergesetzt sind. Weiters fehlen zumeist die kognitiven Operatoren für die Einordnung des Wissens.

Die groupwarebasierte Projektabwicklung liefert daher umfangreichere Informationsquellen als die konventionelle Projektdokumentation, führt aber nur in geringem Maße zu einer gesteigerten Externalisierung von Wissen. Wissensgenerierung kann nur durch eine systematische Externalisierung und Kombination von Wissen erfolgen.

7. Informationsquellenanalyse

Für die Informationsquellenanalyse sind neben der Datenbasis des Groupwaretools auch die folgenden Quellen heranzuziehen. Nur dadurch kann eine ausreichende Menge kognitiver Operatoren zur Einordnung der Informationen zu einem konkreten Projekt geschaffen werden. Die Aufbereitung der im folgenden beschriebenen Informationsquellen muß nicht notwendigerweise gleichzeitig mit der Aufbereitung des Wissens aus einzelnen Projekten erfolgen, sondern wird vielmehr vorgelagert stattfinden müssen und durch einzelne Projekte Modifikationen erfahren.

7.1 Organisationale Routinen

Unter organisationalen Routinen sollen Verfahrensweisen verstanden werden, die zur Lösung bestimmter Aufgaben angewendet werden. Dazu zählen u. a. Entscheidungsprozesse, wie die bereits angesprochenen formellen und informellen Vorgehensweisen zur Entscheidungsfindung in Projekten.

⁷ Zum Begriff der kognitiven Elemente siehe Kapitel 8.

Nach Cyert und March (1963) stellen Standardprozeduren Entscheidungsregeln für bestimmte Situationen dar. Im Rahmen der Informationsquellenanalyse sind daher diese Routinen zu erheben und zu dokumentieren. Diese Externalisierung dieses üblicherweise impliziten Wissens kann durch eine Modellierung⁸ von Vorgehensweisen zur Problemlösung erfolgen. In die Wissensdatenbank wird dann das Modell der jeweiligen organisationalen Routine aufgenommen.

Die dokumentierte organisatorische Routine dient dann als kognitiver Operator, insofern als die im Projekt tatsächlich angewendeten Vorgehensweisen dem Standardmodell gegenübergestellt werden. Weiters erlaubt diese Form der Dokumentation die Einordnung von einzelnen kognitiven Elementen, indem diese den jeweiligen Schritten der Standardroutine zugeordnet werden. Als Beispiel dazu kann ein Zwischenbericht angeführt werden, der im Rahmen des informellen Teiles der Entscheidungsfindung angefertigt wurde. Ohne einer Zuordnung zur entsprechenden Projektphase ist der Inhalt u. U. nur bedingt einzuordnen. Eine detailliertere Darstellung zu allgemeinen organisationalen Routinen findet sich in Pawlowsky (1994, S. 232 ff.).

7.2 Organisationale Handlungstheorie

Auf der individuellen Ebene schafft sich jedes Organisationsmitglied ein Bild von einer Handlungstheorie, die in der Organisation angewendet wird. Der fortlaufende Prozeß der Wirklichkeitskonstruktion erfolgt durch die Kommunikation mit anderen Organisationsmitgliedern. (Argyris und Schön 1978 nach Pawlowsky 1994, S. 235)

Bestandteil einer organisationalen Wissensdatenbank sollte daher auch das Leitbild der Organisation sein, das u. a. Verhaltenserwartungen für bestimmte Situationen und insbesondere die tatsächlich im Gebrauch befindlichen Handlungstheorien – Partizipation bei Entscheidungen – beinhaltet. Eine detailliertere Darstellung zu allgemeinen organisationalen Routinen findet sich in Argyris und Schön. (Argyris und Schön 1978 nach Pawlowsky 1994, S. 236)

Als kognitiver Operator dient dieser Bestandteil der Wissensdatenbank insofern, als die Aufbereitung einen Abgleich zwischen den tatsächlichen und den erwarteten Verhaltensweisen enthalten muß.

⁸ Zum Beispiel durch die Nutzung von Geschäftsprozeßmodellierungstools mit der jeweiligen Methodik.

7.3 Informationssystem

Die Prozesse und Instanzen, die Informationen und Entscheidungen treffen, sind ebenfalls als Bestandteil der Wissensdatenbank aufzufassen, da Individualentscheidungsprozesse nicht isoliert nebeneinander stehen, sondern als Teilprozesse in einem gesamten Informations- und Entscheidungssystem zu sehen sind.

Die Einordnung von Entscheidungen im Rahmen eines Projektes wird daher auch vielfach die Kenntnis des Informationssystems der Organisation erfordern. Somit ist auch dieser Bestandteil der Wissensdatenbank ein wesentlicher kognitiver Operator.

7.4 Störungs-Reaktions-Verknüpfungen

Eine Organisation kann auf jede Störung mit verschiedenen Reaktionen begegnen, die wiederum zu unterschiedlichen Erfolgen führen. Gerade in der Methodik der Projektarbeit wird dem frühzeitigen Erkennen von möglichen Störungen und den denkbaren Reaktionen breiter Raum geschenkt. Umfeldanalysen und Claim-Management sind nur zwei Beispiele dazu. In der Wissensdatenbank sollten daher die in der Organisation üblichen allgemeinen „Störungs-Reaktions-Verknüpfungen“ aufgenommen werden. Diese können dann bei der Aufbereitung der projektspezifischen Ergebnisse als kognitive Operatoren herangezogen werden.

8. Informationsaufbereitung

Im Hinblick auf die Erzeugung informationeller Mehrwerte kann zwischen einem formalen und pragmatischen Ansatz der Informationsaufbereitung unterschieden werden. Während zu den formalen Methoden die mediale Informationsaufbereitung zählt, beschäftigen sich pragmatische Verfahren mit der Anpassung von Informationen an unterschiedliche Benutzerbedürfnisse. Dieser pragmatische Ansatz stützt sich u.a. auf Benutzermodelle und Modelle des organisationellen Zielrahmens. (Kuhlen 1996, S. 88 f.)

Eine Wissensdatenbank, die den Anforderungen einer Lernenden Organisation genügt, kann erst durch eine gezielte Aufbereitung der Informationen, die z.B. die Handlungsrelevanz und den organisatorischen Kontext sicherstellen, gewährleistet werden. (vgl. dazu auch Kuhlen 1996, S. 352 ff.)

Als Ansatz zur Konzeptualisierung der Strukturen und Wirkungsweisen von Wissenssystemen wird das Konzept des Referenzrahmens herangezogen. (Pawlowsky 1994, S. 205-209)

Dieser unterscheidet sechs Kriterien:

Kognitive Elemente

Kognitive Elemente sind die kleinsten Einheiten des Wissenssystems. In der Projektarbeit gehören dazu Dokumente (Projektauftrag, Zielkatalog, etc.), Berichte und alle verwertbaren Bestandteile der Datenbasis des Groupwaretools.

Kognitive Operatoren

Kognitive Operatoren steuern die Organisation von Wissensbestandteilen. Kategoriale Schemata steuern die Einordnung von Informationen, Modellen, Hypothesen etc. In der Projektarbeit sind dies insbesondere die im Abschnitt 7 angeführten Wissensbestandteile.

Realitäts-Tests

Stellen eine Prüfinstanz für kognitive Elemente und Operatoren dar. Vorstellungsinhalte werden validiert, indem Gültigkeitskriterien für die Annahmen herangezogen werden. Unterschieden werden dabei vier idealtypische Realitätstests (Pawlowsky 1994, S. 205): Übereinstimmung von Experten, Konsistenz der Logik als Prinzip, Kongruenz zwischen Theorie und Befunden und Dialektischer Prozeß der Ableitung einer Synthese aus These und Antithese.

Kognitive Landkarten

In der ursprünglichen Konzeption wurden kognitive Landkarten als die kognitiven Bereiche, die für bestimmte Problemstellungen zu aktivieren sind, aufgefaßt. Hier soll eine Adaptierung insofern vorgenommen werden, als kognitive Landkarten für die Methoden des Projektmanagements aufgesetzt werden. Aus der Umfeldanalyse kann beispielsweise aufbereitet werden, welche Personengruppen mit welchen Interessen bei der Projektarbeit zu berücksichtigen sind. Analog dazu lassen sich aus dem Claim-Management die Risikosituation und aus der Informationslandkarte das Projektmarketing erfassen.

Grad der Artikulation

Unter dem Grad der Artikulation versteht man das Ausmaß in dem die Bestandteile der Wissensdatenbank zugänglich sind. Dies erfordert lediglich eine organisatorische Festlegung welcher Personenkreis auf welche Bestandteile zugreifen darf und die Umsetzung mittels Zugriffsrechten.

Metaphern

Unter Methaphern werden Annahmen von Entscheidungsträgern verstanden, die Entscheidungsprozesse steuern. Damit wird angestrebt, in die Wissensdatenbank ein sozial verbindliches Netzwerk von Vorstellungsinhalten einzubinden. Damit korrespondiert dieser Teil sehr stark mit den organisatorischen Handlungstheorien (siehe Abschnitt 7.2). Bei der Aufbereitung des Wissens aus einem konkreten Projekt wären daher hier die

projektspezifischen Vorstellungsinhalte zu erfassen und in Verbindung mit den generellen Aspekten der organisatorischen Handlungstheorien zu setzen.

9. Schlußbetrachtung

Die Aufgaben des Wissensmanagements in Lernenden Organisationen bestehen darin, die organisatorischen und infrastrukturellen Voraussetzungen für den Aufbau einer organisationalen Wissensbasis zu schaffen. Im Kontext eines groupware-unterstützten Projektmanagements zeigt sich, daß die Nutzung und Fortentwicklung einer Wissensdatenbank ein geeignetes Hilfsmittel für Lernende Organisationen darstellt.

Die entwickelten Methoden zum Aufbau einer organisationalen Wissensdatenbank sind insofern erfolgreich, als das der Anteil des expliziten Wissens deutlich erhöht wird.

Der kritische Erfolgsfaktor der Wissensgenerierung liegt jedoch in der Berücksichtigung der Interdependenzen zwischen dem organisatorischen Kontext und der Informationsaufbereitung. Wie gezeigt wurde, sind zur Aufbereitung organisationsrelevanten Wissens genaue Kenntnisse der Organisation bzw. der Projekthinhalte unbedingt notwendig.

Wissensmanagement bietet wichtige Einsatzgebiete für die Informationswissenschaft: InformationswissenschaftlerInnen müssen aber dafür Koalitionen mit Wissensträgern in der Organisation eingehen und ihnen aus dem Informationsparadoxon „drowning in data, yet short of intelligence“ helfen.

10. Literaturverzeichnis

[Argyris C., Schön D.A. 1978]

Organizational Learning. Reading, MA: Addison-Wesley.

[Bullinger H.-J. u.a. 1998]

Produktionsfaktor Wissen. In: Personalwirtschaft, 5/98, S. 22-26.

[Cyert R. M., March J. G. 1963]

A Behavioral Theory of the Firm. Englewood Cliffs, New York: Prentice Hall.

[Kresbach-Gnath C. 1996]

Organisationslernen: Theorie und Praxis der Veränderung. Wiesbaden: Deutscher Universitätsverlag.

[Kuhlen R. 1996]

Informationsmarkt – Chancen und Risiken der Kommerzialisierung von Wissen, 2. Aufl., Konstanz: Universitätsverlag Konstanz.

[Nonaka I., Takeuchi H. 1997]

Die Organisation des Wissens: wie japanische Unternehmen eine brachliegende Ressource nutzbar machen. Frankfurt: Campus Verlag.

[Österle H., Vogler P. 1996]

Praxis des Workflow-Managements. Grundlagen, Vorgehen, Beispiele. Braunschweig: Vieweg.

[Ortwein E., Spallek P. 1998]

Wissensmanagement in der Managementberatung – Praxisbeispiel Arthur D. Little. In: *IM – Die Fachzeitschrift für Information Management & Consulting*, 1/98, S. 105-107.

[Pawlowsky P. 1994]

Wissensmanagement in der lernenden Organisation. Unveröffentlichte Habilitationsschrift. Paderborn.

[Pawlowsky P., Reinhardt R. 1997]

Wissensmanagement: Ein integrativer Ansatz zur Gestaltung organisationaler Lernprozesse. In: *Handbuch Lernende Organisation: Unternehmens- und Mitarbeiterpotentiale erfolgreich erschliessen*. Wiesbaden: Gabler.

[Penker W., Proske D. 1993]

Notenbankforschung als Wettbewerbsfaktor in einer Wirtschafts- und Währungsunion. In: *Wirtschaftspolitische Blätter*, 6/93, S. 680-693.

[Picot A., u.a. 1998]

Die grenzenlose Unternehmung. Information, Organisation und Management, 3. Aufl., Wiesbaden: Gabler.

[Probst G. 1997]

Wissen managen. Wie Unternehmen ihre wertvollste Ressource optimal nutzen. Wiesbaden: Gabler.

[Probst G., Büchel B. 1998]

Organisationales Lernen. Wettbewerbsvorteil der Zukunft, 2. Aufl., Wiesbaden: Gabler.

[Rey M. u.a. 1998]

Studenmodell zur Einführung von Wissensmanagement. In: *Information Management*, 1/98, S. 30-36.

[Senge P. 1990]

The Fifth Discipline: The Art and Practice of the Learning Organisation. New York: Doubleday.

[Schuller F. 1998]

Wissensaufbau erfordert eine offene Lernkultur. In: Personalwirtschaft, 5/98, S. 27-30.

[Servatius H.-G. 1998]

Intellektuelle Wertschöpfung mit Wissensnetzwerken. In: IM - Die Fachzeitschrift für Information Management & Consulting, 1/98, S. 101-104.

[Wagner M.P. 1995]

Groupware und neues Management. Einsatz geeigneter Softwaresysteme für flexiblere Organisationen. Braunschweig: Vieweg.

Section 2

Multimedia and Knowledge Transfer

(Electronic Market, Knowledge Exchange)

Bedarf, Architektur und Nutzen elektronischer Marktplätze - einige Konstruktionshinweise dargestellt am Beispiel des Marktplatzes Bildung

JOSEF HERGET, SILKE GROSSMANN

Dr. Josef Herget,
Silke Grossmann

IMAC Information & Management Consulting, Konstanz

Blarerstr. 56, D-78462 Konstanz
josef.herget@imac.de

- 1 KONTEXT UND KONZEPT DER STUDIE
- 2 FACHINFORMATION IM BILDUNGSWESEN
 - 2.1 DIE BEDEUTUNG VON BILDUNGS(FACH)INFORMATION
 - 2.2 ZUR SITUATION BILDUNGSBEZOGENER FACHINFORMATION
- 3 FACHINFORMATION AUF ELEKTRONISCHEN MÄRKTEN
 - 3.1 GESTALTUNGSOPTIONEN ELEKTRONISCHER FACHMÄRKTE
 - 3.2 MODELL EINES ELEKTRONISCHEN MARKTPLATZES FÜR BILDUNG
- 4 ZIELGRUPPENANAYSE
- 5 INFORMATIONSBEDARFSANALYSE
- 6 STRATEGISCHE MARKETINGMAßNAHMEN FÜR EINEN ELEKTRONISCHEN MARKTPLATZ BILDUNG
- 7 MÖGLICHE ORGANISATIONS- UND BETREIBERMODELLE
- 8 AUFGABEN FÜR DEN AUFBAU UND BETRIEB EINES ELEKTRONISCHEN MARKTES
- 9 LOGISCHES DESIGN UND REALER PROTOTYP
- 10 ZUSAMMENFASSUNG UND AUSBLICK

ZUSAMMENFASSUNG

Neue Medien und die verfügbare Kommunikationsinfrastruktur ermöglichen neue Gestaltungsoptionen für Informationsanbieter. In diesem Beitrag wird ein Vorgehenskonzept am Beispiel der Bildungsinformation beschrieben. Ausgehend von der Bestimmung relevanter Zielgruppen wird deren Informationsbedarf erhoben und als Anforderung für neue Informationsdienste umgesetzt. Hinweise zur Ausgestaltung der Marketingmaßnahmen und zur Architektur elektronischer Marktplätze werden am konkreten Beispiel des Elektronischen Marktplatzes Bildung erläutert. Die Darstellung eines Prototyps schließt die Ausführungen ab.

1 KONTEXT UND KONZEPT DER STUDIE

Telematik und Neue Medien verändern im erheblichen Maße die Informations- und Kommunikationsinfrastruktur und schaffen damit zugleich zahlreiche Potentiale zur effektiven und effizienten Interaktion in Fachgemeinschaften. Eine konkrete Ausprägung dieser veränderten Informationslandschaft stellen *elektronische Marktplätze* dar, die auf den Mehrwertdiensten des Internets beruhen. Fachlich orientierte elektronische Marktplätze als mögliche Geschäftsmodelle auch "klassischer" Fachinformationsanbieter sollen dabei den mit der dezentralen Struktur des Internets für den Informationssuchenden verbundenen Nachteil der Intransparenz vorhandener Angebote ausgleichen und virtuelle Informationsräume schaffen, in denen alle bei der Informationssuche relevanten Prozesse unterstützt werden.

In der Zeit von Dezember 1997 - April 1998 wurde für das Deutsche Institut für Internationale Pädagogische Forschung (DIPF) in Frankfurt a. M. eine empirisch gestützte Studie angefertigt. Ziel war es, ausgehend von der aktuellen Situation im Bereich der Bildungsfachinformation, Nutzererwartungen an einen fachorientierten elektronischen Marktplatz zu erarbeiten und die Definition eines adäquaten Leistungsangebotes für einen "Elektronischen Marktplatz Bildung" vorzunehmen.

Im Rahmen der Studie wurde folgenden Fragestellungen nachgegangen:

- ◆ Erarbeitung des State-of-the-Art im Hinblick auf die Gestaltungsoptionen elektronischer (Fach)Märkte
- ◆ Identifizierung der potentiellen Zielgruppen eines elektronischen Marktplatzes Bildung und deren Informationsbedarf
- ◆ Festlegung strategischer und operativer Maßnahmen bei der Realisierung des Marktplatzes
- ◆ Bestimmung möglicher Organisations- und Betreibermodelle und deren Stärken und Schwächen.

Um die Beurteilung der Funktionalität und Attraktivität zu erleichtern, wurde auf der Grundlage der durchgeführten Analysen und der erarbeiteten Ergebnisse ein Prototyp entwickelt, der einer ersten Evaluation dienen sollte.

Für die Erreichung dieser Ziele kamen folgende Methoden zur Anwendung:

- ◆ Benchmarking bestehender Angebote im Bildungsbereich und anderer Branchen
- ◆ Interviews mit Vertretern der potentiellen Zielgruppen
- ◆ Konzeptuelle Überlegungen zu den Marketingmöglichkeiten im WWW
- ◆ Literaturanalyse und Fallstudien
- ◆ Entwurf, Layout, Gestaltung und Programmierung des Marktplatzes auf einem WWW-Server

Das für diese Studie gewählte Vorgehen lässt sich in folgenden aufeinander aufbauenden Phasen beschreiben:

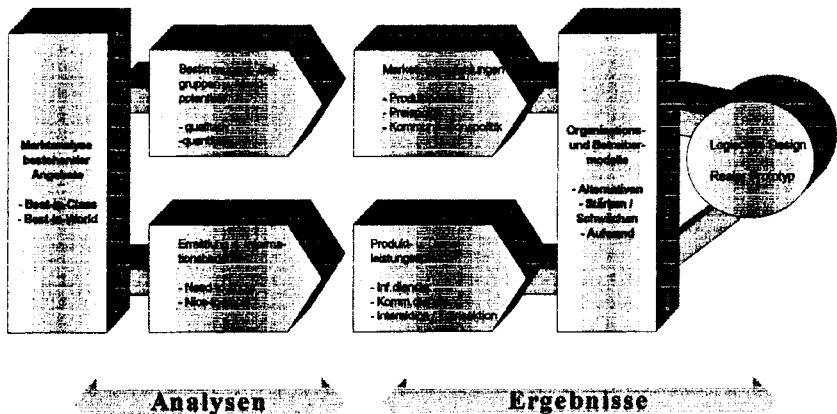


Abb. 1: Vorgehenskonzept der Studie

2 FACHINFORMATION IM BILDUNGSWESEN

2.1 DIE BEDEUTUNG VON BILDUNGS(FACH)INFORMATION

In einer Zeit, in der "Wissen als wichtigster Rohstoff" angesehen wird, welcher "nur durch Bildung zu erschließen ist" (Herzog 1997), in der der Funktionswandel des Bildungssystems als notwendige Voraussetzung für die Vorbereitung auf die

Informationsgesellschaft angesehen wird und in der Schlagworte wie "Lebenslanges Lernen" und "Recurrent Education" und die durch die neuen Informations- und Kommunikationstechnologien eingeleitete Umgestaltung aller Bereiche des Beschäftigungssystems neue Lern- und Qualifikationsanforderungen begründen, wird eine informationelle Unterstützung aller Betroffenen immer wichtiger werden. Die Explosion der Bildungsangebote und -produkte vor allem im Weiterbildungsbereich, die starke Differenzierung des Bildungswesens und der Bildungswege erschwert zusätzlich, insbesondere dem privaten Bildungsnachfrager, die Orientierung auf und über den "Markt". Bildungs(fach)information gewinnt in diesem Zusammenhang einen Bedeutungszuwachs, welcher zum einen ein großes Potential (i. S. von Bedarf) für die Zukunft begründet, zum anderen aber sicherlich auch ein Umdenken in der Frage, was Fachinformationstätigkeit nun eigentlich umfassen und an wen sie sich richten sollte, erfordert.

2.2 ZUR SITUATION BILDUNGSBEZOGENER FACHINFORMATION

Am Fachinformationsmarkt des Bildungswesens sind eine große Anzahl unterschiedlicher Teilnehmer beteiligt: Rundfunk- und Fernsehanstalten, Landesmedienstellen bzw. -institute, Verlage, andere kommerzielle Medienanbieter und die Bildungseinrichtungen selber. Das Deutsche Institut für Internationale Pädagogische Forschung beherbergt mit dem Fachinformationssystem Bildung (FIS Bildung) die zentrale Fachinformationseinrichtung für das Bildungswesen in Deutschland. Die mit dem FIS Bildung kooperierenden Einrichtungen erstellen sowohl selbst als auch im Verbund des FIS-Bildung Informationsdienstleistungen, die im wesentlichen auf der Literatur des Fachgebietes beruhen.

Eine 1992 durchgeführte Bestandsaufnahme beim FIS Bildung nennt als wichtigste Nutzergruppen der angebotenen Informationsdienstleistungen Wissenschaft / Forschung (28%) und die pädagogische Praxis (24%). Eine große Rolle spielen auch Mitarbeiter und Gäste der Informationsdienstleister (12%), Bildungsverwaltung/ -politik (10%) und die interessierte Öffentlichkeit (10%). Darunter wird die Fachöffentlichkeit subsumiert, die den dargestellten Kategorien nicht zuzuordnen ist. Es folgen noch Informations- und Dokumentationseinrichtungen (5%), privatwirtschaftliche Einrichtungen (3%), Verbände / Gewerkschaften (3%), Journalisten (2%), Medizin / Psychologie (1%), Eltern (1%) und Sonstige (1%). Allerdings läßt diese Untersuchung keine Aussagen zu, wie hoch der Anteil der einzelnen Gruppen an der tatsächlichen Nutzung ist. Die Verteilung der tatsächlichen Nutzung ist aber nach Aussage der im Rahmen der Studie 52 befragten Institutionen sehr ungleich, mit der Hauptnutzergruppe Wissenschaft und Forschung (inklusive Studenten) - als "klassische Klientel" der durch die Literaturdokumentation geprägten Fachinformation.

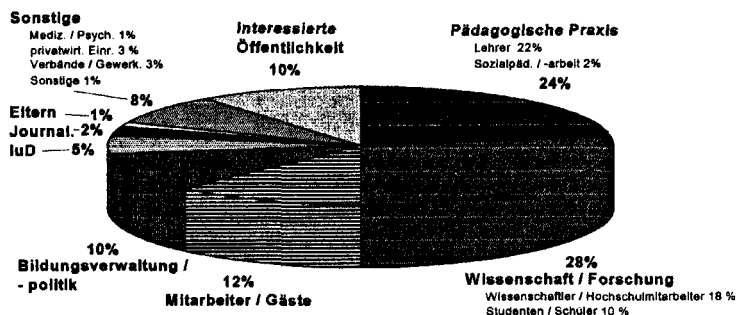


Abb. 2: Nutzerklassen bestehender Fachinformationsleistungen im

3 FACHINFORMATION AUF ELEKTRONISCHEN MÄRKTEN

3.1 GESTALTUNGSOPTIONEN ELEKTRONISCHER FACHMÄRKTE

Die durch verschiedene Benutzerstudien schon relativ früh gewonnene Einsicht, daß der Bedarf nach Daten und Fakten wächst, während der nach Literaturinformation zurückgeht (Kunz/Rittel/Schwuchow 1977: 60), hat in vielen Bereichen zur veränderten Betrachtung des Fachinformationsbegriffes geführt. Das bringt, bedingt auch durch die rasante Entwicklung des Informations- und Präsentationsmediums WWW, eine neue veränderte Ausgangslage für die Fachinformationstätigkeit, die sich folgendmaßen charakterisieren läßt:

- ◆ Erschließung neuer Zielgruppen, die auch für den Fachinformationsmarkt und seine Leistungen erreichbar werden
- ◆ Generierung neuer Informationsprodukte, die erst durch die spezifischen Möglichkeiten des Internets entstehen können
- ◆ zunehmende Verlagerung der Fachkommunikation bzw. des fachlichen Austausches in elektronische Netze (Kuhlen / Zhang 1997).

Diese Veränderungen begründen den höheren Stellenwert des Mediums "Online" und letztlich die Potentiale des Internets für den klassischen Fachinformationsmarkt. Leitet man daraus die Notwendigkeit des Aufbaus und Angebotes neuer Informationsdienste ab, so wird der elektronische Fachmarkt zum möglichen Modell der Informationsarbeit im und mit dem Internet.

Vor allem drei Aspekte sind von hoher Bedeutung für den Aufbau elektronischer Fachmärkte:

- ◆ die Qualität der Informationsangebotes, die durch den Betreiber des Fachmarktes garantiert werden muß. Die durch das wachsende Informationsangebot vor allem unter qualitativen Gesichtspunkten erschwerte Informationsselektion wird dabei für den Nutzer immer schwieriger alleine zu bewältigen sein. Der Betreiber sollte in dieser Situation unterstützend wirken. Die Relevanzeinschätzung kann durch detaillierte inhaltliche Beschreibungen der Informationsquelle erleichtert werden, das Vertrauen zu einer Informationsquelle kann erweckt werden, indem die qualitativen Merkmale einer Informationsquelle auf dem Fachmarkt transparent gemacht werden. Nur wenn der potentielle Nutzer den Fachmarkt mit bestimmten Qualitätseigenschaften in Verbindung bringen kann, wird er diesen anderen Wegen der Informationsbeschaffung vorziehen und vielleicht auch bereit sein, für diesen Mehrwert zu bezahlen.
- ◆ der Aufbau von Kommunikationsnetzwerken. Neben den formellen Kanälen der Wissenschaftskommunikation (Zeitschriften, Monographien, Bibliographien) spielen in der Fachkommunikation auch informelle Kanäle zum Informationsaustausch sogenannte Beziehungsnetzwerke eine wichtige Rolle. Den Aufbau dieser Kommunikationsstrukturen zur Unterstützung des informellen Austausches unter Ausnutzung der Möglichkeiten der Mehrwertdienste des Internets zu unterstützen und zu fördern, stellt einen weiteren wichtigen Erfolgsfaktor für elektronische Fachmärkte dar.
- ◆ die Öffnung des Fachmarktes in die Publikumsmärkte hinein. Das Angebot auf den Fachmärkten muß dahingehend erweitert werden, daß auch der interessierte Laie in die Lage versetzt wird, sein Bedürfnis nach Fachinformation zu befriedigen.

Diese Gestaltungsanforderungen an elektronische Fachmärkte im Internet müssen durch die Forderung nach einer Erbringung von Orientierungs- und Strukturierungsleistungen ergänzt werden. Elektronische Märkte sind Mittler(märkte), deren Leistungen "... in erster Linie Orientierungs- oder allgemeiner: Strukturierungsleistungen ..." sind (Kuhlen 1996:19). Wesentlich ist die Schaffung eines (virtuellen) Informationsraumes, in dem alle bei der Informationssuche relevanten Prozesse unterstützt werden. Wo sich die benötigten Informationen dabei befinden bzw. wer diese anbietet ist nicht ausschlaggebend, der Benutzer muß nur wissen, daß es sie gibt, unter welchen Bedingungen er darauf zugreifen kann und letztlich auch den unmittelbaren Zugriff über den elektronischen Markt realisieren können.

3.2 MODELL EINES ELEKTRONISCHEN MARKTPLATZES FÜR BILDUNG

Ohne bereits die Bedürfnisse der potentieller Zielgruppen eines Elektronischen Marktplatzes Bildung zu berücksichtigen, lassen sich die diskutierten Anforderungen an Gegenstand und Funktionen eines Fachmarktes in einem Modell zusammenfassen, das in Abbildung 3 dargestellt ist. Im Mittelpunkt stehen die Informationen, auf die sich der Bedarf potentieller Nutzer richtet ("Bildungsinformationen") - theoretisch all die Informationen, die im Zusammenhang mit dem Bildungs- und Forschungsprozeß und Verwaltungs- und Planungsprozeß

von Bildungsangeboten stehen und entstehen. Der elektronische Marktplatz wird zu einem Informationssystem, das die mit diesen Informationen in Verbindung stehenden Angebote zusammenführt und diese transparent macht. Neben dieser Informationsfunktion des Fachmarktes muß die Transaktionsfunktion treten, damit der Nutzer auch die Informationen bekommt, die er benötigt. Am Abschluß der Transaktionsphase steht also die Übermittlung des angeforderten Gutes selber (Schmid / Lindemann 1997: 4-7). In diesen Transaktionsphasen, die auf unterschiedliche Weise unterstützt werden können, liegt das größte Potential für den Marktplatzbetreiber, denn hier kann er Mehrwerte schaffen und dafür sorgen, daß der Besucher auch nach Inanspruchnahme der Vermittlungsleistung immer wieder kommt (Stichwort: SDI-Dienste, Alert-Dienste, Push-Konzepte, Auftragsrecherchen). Information und Transaktion enthalten als wesentliche Marktfunktionen immer auch kommunikative und interaktive Aspekte. Der elektronische Marktplatz sollte jedoch darüber hinaus auch Kommunikationsplattform für seine Teilnehmer sein, deren vorrangiges Ziel es ist, den fachlichen Austausch zu fördern (durch Mailinglisten, Kommunikationsforen, Expertenbörsen usw.).

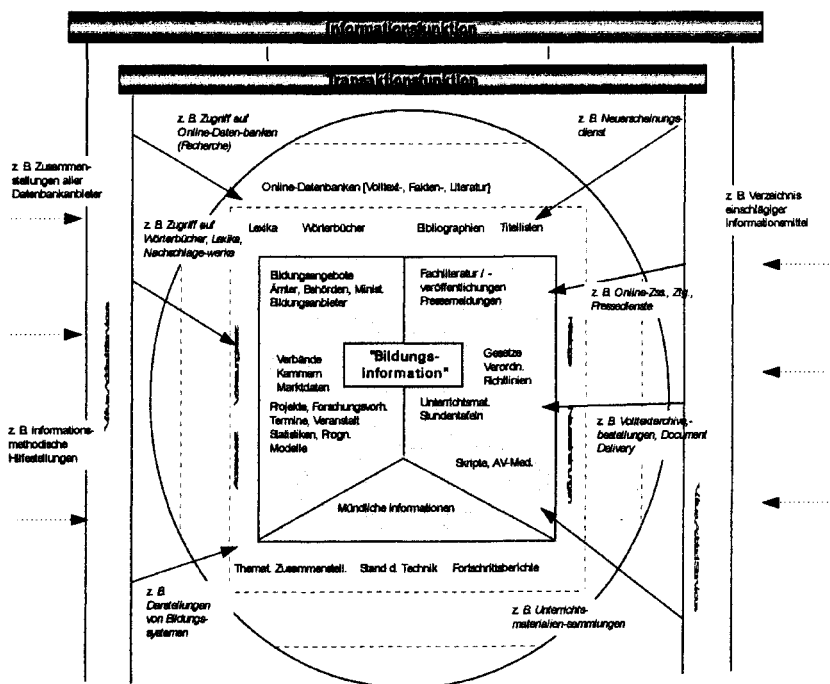


Abb. 3: Gegenstand und Funktionen eines "Elektronischen Marktes Bildung"

4 ZIELGRUPPENANALYSE

Das mit dem Aufbau eines elektronischen Fachmarktes durch des DIPF verfolgte Ziel, das bisherige Leistungs- und Aufgabenspektrum zu erweitern, impliziert, daß ein weitaus größerer Abnehmer- und Kundenkreis ins Blickfeld rücken sollte, als es bisher mit dem vorwiegend auf die Bildungswissenschaft ausgerichteten Angebot der Fall war. Neben den bisherigen Kunden des DIPF, die auch zukünftig als Hauptzielgruppe eines Marktplatzes Bildung zu berücksichtigen sind, werden auch neue Zielgruppen relevant.

Folgende Zielgruppen wurden bestimmt:

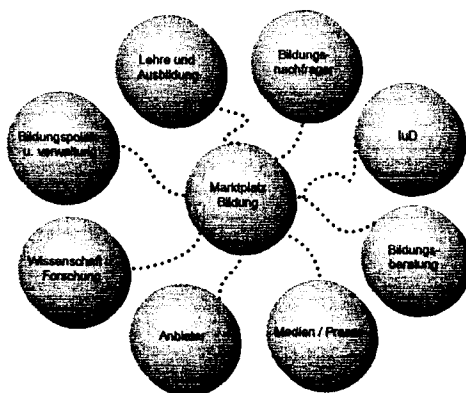


Abb. 4: Zielgruppen des Elektronischen Marktplatzes Bildung

5 INFORMATIONSBEDARFSANALYSE

Die Informationsbedarfsanalyse spielt im Planungsprozeß eines Informationsdienstes eine zentrale Rolle, denn nur wenn Klarheit besteht, welche Zielgruppen welchen Informationsbedarf haben, kann ein elektronischer Marktplatz nutzergerecht konzipiert und aufgebaut werden. Zielsetzung der Analyse war es daher, den Informationsbedarf in den einzelnen pädagogischen Handlungsfeldern zu ermitteln, um daraus Ansatzpunkte für eine Konkretisierung der Inhalte des elektronischen Marktplatzes ableiten zu können. Dabei wurde ein qualitatives Vorgehen gewählt, welches für eine erste Annäherung an die Problematik besser geeignet erschien als quantitative Methoden. Insgesamt wurden 29 Interviews mit Personen durchgeführt, von denen anzunehmen war, daß sie über eine fundierte und verlässliche Meinung bezüglich des Untersuchungsgegenstandes verfügen (Experten). Gegenstand der Befragung waren dabei sowohl die unabhängig vom Internet für die Aufgabenerfüllung benötigten Informationen, als auch die speziell

auf das Internet bzw. einen potentiellen Marktplatz Bildung gerichteten Nutzenerwartungen und Leistungsanforderungen. In einem weiteren Schritt wurde den Befragten ein Fragebogen zugesandt, in dem sie zu den systematisierten Informationsbedarfsäußerungen der Partizipanten aus den Interviews befragt wurden.

Die Auswertung der Interviews und des Fragebogens ergab sehr vielfältige Hinweise auf die Bedürfnisse und Erwartungen der Befragten in Hinblick auf einen Elektronischen Marktplatz Bildung. Die Ergebnisse lassen sich unabhängig von einer zielgruppenspezifischen Betrachtung dahingehend systematisieren, daß es Informationsarten und Themen gibt, die für alle Interessensgruppen von Bedeutung sind:

- | | |
|-----------------------------|-------------------------------------|
| ■ Personen / Experten | ■ Programme und Projekte |
| ■ Thematische Literaturinf. | ■ Informationen mit aktuellem Bezug |
| ■ Adressen / Institutionen | ■ Informationsmethodik / -mittel |
| ■ Daten und Fakten | ■ Termine und Veranstaltungen |

Daneben wird bei einer Realisierung eines Marktplatzes zu berücksichtigen sein, daß einige Dienste für einen großen Teil der Zielgruppen interessant sind und wiederum andere möglicherweise nur bestimmten Zielgruppen angeboten werden sollten:

Für Viele interessant:

- Bildungstatistik
- Bildungsrecht
- Literaturinformationen
- Bildungsangebote

Für Wenige interessant:

- Didaktische Materialien
- Bildungsorganisation / -verwaltung
- Bildungsforschung

6 STRATEGISCHE MARKETINGMAßNAHMEN FÜR EINEN ELEKTRONISCHEN MARKTPLATZ BILDUNG

Auf der Grundlage der Ergebnisse der durchgeführten Analysen wurden erste Empfehlungen für das strategische Marketing eines elektronischen Marktplatzes Bildung formuliert.

Produktstrategie:

Die bisherigen Produkte des DIPF sollten die Basis für den Elektronischen Marktplatz Bildung darstellen. Die Bereitstellung und der Zugriff auf diese im WWW kann zur Erhöhung einer Marktdurchdringung beitragen, die gleichzeitig mit einem verhältnismäßig geringen Ressourceneinsatz verbunden sein wird. Voraussetzung ist es allerdings, mit diesem Angebot leistungs- bzw. produktverbessernde

Maßnahmen, sowie Maßnahmen, die auf eine stärkere Kundenbindung abzielen, zu verbinden:

- ◆ Aufnahme von eigentlich vom Kunden zu erbringenden Leistungen (Internalisierung)
- ◆ Individualisierung des Informationsangebotes auf der Grundlage eines effektiven Database-Marketings
- ◆ Integration von Interaktions- und Kommunikationsangeboten, die dem Kunden die Kontaktaufnahme zum Anbieter ermöglichen

Der Bedarf an Literaturinformation ist nach Meinung der Experten vorhanden, die Nutzung der Produkte des DIPF, die in erster Linie ja Literaturinformationsmittel sind, im bzw. über das WWW, erfordert jedoch eine Modifikation dieser:

- ◆ durch die Bereitstellung benutzerfreundlicher Retrievaloberflächen
- ◆ durch die Bereitstellung von Orientierungs- und Navigationsangeboten für den rechercheunerfahrenen Endnutzer
- ◆ durch die Entwicklung neuer Angebotsformen für die bestehenden Produkte

Nur so wird es möglich sein, mit den bestehenden Produkten neue Zielgruppen anzusprechen. Die Entwicklung neuer Produkte für bestehende und neue Zielgruppen rundet die auf die Produktpolitik abzielende Marketingempfehlung dieser Studie ab und soll als Diversifikationsstrategie bezeichnet werden. Neue Produkte im WWW sind weniger neu, sondern stellen vielmehr den Versuch dar, durch andere Aufbereitungsverfahren Mehrwerte zu schaffen und die bisherigen Produkte und Leistungen um zusätzliche Komponenten zu erweitern.

Diese Strategien sollten zur Umsetzung der One-Stop-Shop-Idee führen, realisiert durch den aufzubauenden Marktplatz Bildung. Dieser sollte so konzipiert sein, daß er in der Lage ist, dem Nutzer alle Informationen zu bieten, die dieser zur Erfüllung seiner beruflichen Aufgaben - oder allgemeiner: zur Bewältigung seiner Informationsprobleme benötigt.

Preisstrategie:

Die Produkte und Leistungen des DIPF wurden bisher kostenlos bzw. zu vergleichsweise geringen Preisen angeboten. Für den Marktplatz Bildung wird daher empfohlen, an der Niedrigpreispolitik festzuhalten. Vorgeschlagen wird die Unterscheidung zwischen Basisinformationen, die von jedem Benutzer (kostenlos) abgerufen werden können, und kostenpflichtigen Informationsangeboten. Die Inanspruchnahme kostenpflichtiger Leistungen wird an eine Registrierung des Benutzers gebunden sein. Die Preisgestaltung richtet sich nach dem Umfang der in Anspruch genommenen Leistungen (Pricing-for-Information). Zusätzlich können für die hochpreisigen Informationen Pauschalgebühren eingeführt werden, die unterschiedliche Servicelevel (z. B. Beschränkung der Nutzung auf bestimmte

Datenbanken, Ausschluß der Dokumentenlieferung) und unterschiedliche Nutzerzahlen berücksichtigen. Der Nutzer hat somit die Möglichkeit, die Kosten der Marktplatznutzung individuell zu steuern und das auf seine Bedürfnisse am besten zugeschnittene Preismodell auszuwählen.

Kommunikationsstrategie:

Bestandteil einer effektiven Kommunikationsstrategie für einen internetbasierten Dienst muß immer auch die traditionelle Pressearbeit sein - zumal die Experteninterviews deutlich gemacht haben, daß der Öffentlichkeitsarbeit in der Vergangenheit nicht die ihr gebührende Bedeutung zugemessen wurde. Für eine Vermarktung des Marktplatzes ist daher zwingend ein offensivere Kommunikationspolitik zu empfehlen, die Elemente der klassischen Kommunikationspolitik und Mittel, die im WWW zur Verfügung stehen, verbindet (integrierte Kommunikationsstrategie).

7 MÖGLICHE ORGANISATIONS- UND BETREIBERMODELLE

Elektronische Märkte können bezüglich der Organisation und des Betriebes in verschiedenen alternativen Formen gestaltet werden. Ein elektronischer Marktplatz besteht aus folgenden konzeptuellen Ebenen, auf denen verschiedene Aufgaben und Tätigkeiten anfallen, die Auswirkungen auf die Wahl des Betreibermodells haben können:

- ◆ Informationsebene
- ◆ Generische Ebene
- ◆ Technische Ebene

Die Realisierungsmöglichkeiten für einen Marktplatz Bildung auf technischer Ebene lassen sich wie folgt unterscheiden:

- ◆ Zentraler Serverbetrieb
- ◆ Dezentraler Serverbetrieb
- ◆ Verteilter Serverbetrieb

Hinsichtlich der Informations- und Generikebene lassen sich folgende Aussagen treffen:

- ◆ Das Angebot und die notwendigen Dienste der Informationsebene, die als Ergebnis der Bedarfsanalyse bestimmt werden konnten, erfolgen unabhängig von der Art der physikalischen Serverstandorte und können über das Internet für jede Form der Serverkonstellation realisiert werden, da die Dienste der Informationsebene marktplatzbezogen zu sehen sind. Häufig übernimmt diese

nicht der Provider. Auch hier ist zwischen zentralem, dezentralem und verteiltem Serverbetrieb zu unterscheiden.

- ◆ Für das Angebot der Generikdienste - dazu zu zählen sind z. B. intelligente Suchassistenten, Newsservice, Zahlungsverkehr, Werbung - gilt analog zu den Informationsdiensten: Dienste der Generikebene sind unabhängig von der Art der physikalischen Serverstandorte und können über das Internet für jede Form der Serverkonstellation realisiert werden. Häufig übernimmt diese der Provider, der über das notwendige technologische Know-How verfügt. Die Entwicklung der Dienste auf der Informations- bzw. Generikebene erfolgt in der Regel in Zusammenarbeit von Provider und Marktplatzanbieter.

8 AUFGABEN FÜR DEN AUFBAU UND BETRIEB EINES ELEKTRONISCHEN MARKTES

Zu den Aufgaben, die mit der technischen Betreuung des Marktplatzes Bildung zusammenhängen, zählen:

Für die Technikenebene:

- ◆ der Serverbetrieb (Einrichtung, laufende Überwachung und Monitoring)
- ◆ Datenschutz und -sicherheit, Backup
- ◆ Anbindung des Servers an das Internet (Access-Providing)
- ◆ Erstellung von Abruf- und Zugriffsstatistiken (quantitative und qualitative Erfolgskontrolle)

Für die Generikebene:

- ◆ die Programmierung von Schnittstellen für marktplatzübergreifende Mehrwertfunktionen
- ◆ Entwicklung weiterer Aspekte zur Steigerung des Zusatznutzens
- ◆ Suche nach Schlagwort und/oder Thesaurus
- ◆ Intelligente Suchfunktionen
- ◆ Zahlungssysteme
- ◆ elektronische Abwicklung der Geschäftsprozesse
- ◆ Realisierung der Kommunikationsdienste

Die Aufgaben für die inhaltliche Betreuung des Marktplatzes beziehen sich auf den Aufbau und die Weiterentwicklung des Angebotes und sind konzeptioneller und

redaktioneller Natur. Sie fallen für die Bereitstellung der Dienste der Informations- und Generikebene an. Diese Aufgaben werden dabei sehr stark vom jeweiligen Produkt- und Dienstleistungsspektrum des Marktplatzes abhängen, so daß der nötige Aufwand nur beispielhaft abgesteckt werden kann.

Aufgaben der Informationsebene:

- ◆ Design und Gestaltung des Marktplatzes
- ◆ Kontrolle und Verwaltung der Marktplatzeinträge (Name, Beschreibung, URL) mit regelmäßiger Prüfung auf URL-Existenz, z. B. Links auf Literaturinformation/ Literaturdokumentation, Dokumentation von Lehr- und Lernmitteln, Fakteninformation (laufende Qualitätssicherung)
- ◆ redaktionelle Betreuung von inhaltlichen Beiträgen (laufend)
- ◆ Bereitstellung der Informationsangebote des Marktplatzes

Aufgaben der Generikebene:

- ◆ Intelligente Suchagenten
- ◆ Pflege von Interessensprofilen
- ◆ Kommunikationsdienste
- ◆ Schaltung / Anfertigung von Minibannern, evtl. durch Ad-Server unterstützt
- ◆ Schaltung von Links an prominenter Stelle
- ◆ Abrechnung von Gebühren
- ◆ zielgruppenspezifische Werbeeinblendungen (Keyword Advertising), Akquisition von Werbung
- ◆ Betreuung der Mitglieder für geschlossene Bereiche (mit oder ohne Abrechnungsfunktionen)
- ◆ Verkaufsförderung, Cross-Media Publishing, eigene Vermarktung im Internet und Erfolgskontrolle etc.

Je nach Ausbaustufe des Marktplatzes können weitere Aufgaben hinzukommen. Es empfiehlt sich, weite Teile dieses Aufgabenbereiches durch Verwaltungssysteme zu unterstützen (z.B. dynamische Generierung der Seiten, Ad-Server für Werbeverwaltung, Email-Verteilerlisten, Majordomo-Listserver). Nicht alle Aufgaben werden dabei vom Marktplatzbetreiber wahrgenommen werden müssen, sondern können an andere Dienstleister vergeben werden.

9 LOGISCHES DESIGN UND REALER PROTOTYP

Im folgenden soll ein Produkt- und Dienstleistungsportfolio für einen elektronischen Marktplatz Bildung zusammengestellt werden. Ziel sollte es sein, die auf das WWW bezogenen Marketingempfehlungen und die in Anlehnung an die Ergebnisse der Marktanalyse vorgeschlagenen Strategien geeignet umzusetzen. Besondere Berücksichtigung sollten dabei folgende Ansatzpunkte für eine Differenzierung von bestehenden Angeboten finden:

- ◆ Realisierung eines breiten Informationsangebotes
- ◆ Ausstattung der bestehenden Produkte mit Zusatznutzen
- ◆ hoher Grad an Interaktion
- ◆ Direkter Zugriff auf die Produkte und Dienstleistungen durch verschiedene Transaktionsmodi

Beim Entwurf des Produkt- und Dienstleistungsportfolios wurde in Anlehnung an die Grundfunktionen elektronischer Märkte zwischen Informationsdiensten (Inhalten) einerseits und Interaktions-, Kommunikations- und Transaktionsdiensten andererseits unterschieden. Für die Festlegung der Inhalte des Marktplatzes wurden die Ergebnisse der Informationsbedarfsanalyse berücksichtigt. Unter der Annahme, daß der Marktplatz zunächst die Bedürfnisse aller potentiellen Zielgruppen berücksichtigen sollte, wurde jedoch keine Priorisierung bzw. zielgruppenspezifische Auswahl bereitzustellender Informationen vorgenommen.

Die Informationsangebote des Marktplatzes Bildung können unterteilt werden in Informationen, die der Realisierung kommunikationspolitischer Ziele dienen und bildungsbezogenen Informationen, die unmittelbar auf die Befriedigung des fachlichen Informationsbedarfs der Zielgruppen ausgerichtet sind.

Unternehmensbezogene Informationsangebote:

- ◆ Ausführliche Unternehmensinformationen: "Wir über uns" mit Forschungsschwerpunkten, Mitarbeiter(porträts) / Ansprechpartnern. Im Sinne einer Interaktion / Kommunikation zusätzlich durch Kontaktmöglichkeiten zu diesen: "Mitarbeiter als Experten"
- ◆ Einbindung der Kooperationspartner mit Verzweigungsmöglichkeiten zu den jeweiligen Anbieterseiten (günstigenfalls in Framestruktur, um Verbindung zum Marktplatz zu gewährleisten)
- ◆ Produktinformationen: Veröffentlichungen des DIPF / FIS Bildung und auf den Marktplatz bezogene Informationen (Hilfefunktion), Produkt-schulung in Form von Online-Tutorials
- ◆ "What`s New": Produktneuigkeiten, neue Ressourcen oder sonstige Veränderungen auf dem Server

- ◆ produkt- bzw. unternehmensbezogene Mailingliste
- ◆ Bereitstellung von Materialien für die Presse (Presseservice)
- ◆ FAQs als Interaktionsmöglichkeit und spezifische Form der Produkterklärung
- ◆ Testimonials

Bildungsbezogene Informationsangebote:

Die bildungsbezogenen Informationsangebote des Marktplatzes umfassen die Produkte / Dienstleistungen des DIPF einerseits, neue Informationsdienste andererseits, die den Marktplatz mit Zusatznutzen für den Besucher ausstatten.

Bisherige Produkte / Dienstleistungen des DIPF

- ◆ Recherche in den Datenbanken
- ◆ Möglichkeit, Informationsrecherchen in Auftrag zu geben
- ◆ Neue Angebotsformen für die Produkte
 - ◆ Thematische Zusammenstellung von Literaturhinweisen aus den Datenbanken
 - ◆ Literaturtips / Rezensionen, Neuerscheinungen
 - ◆ Table-of-Contents-Dienst: Blättern in den Inhaltsverzeichnissen der Zeitschriften, die für die Datenbanken ausgewertet werden [soweit diese elektronisch vorliegen oder elektronifiziert werden können]
 - ◆ Zeitschriftendokumentation Bildung: zusätzlich thematische Aufbereitung ausgewählter Beiträge

Neue / zusätzliche Informationsdienste

Neue bzw. zusätzliche Informationsdienste sollen die bestehenden Produkte und Dienstleistungen des DIPF ergänzen und die zu konzipierende Anwendung im Sinne der One-Stop-Shop-Idee zu einer Anlaufstelle für alle an Bildungsinformation Interessierten machen. Erreicht werden kann dieses durch:

- ◆ die Aufbereitung von bestimmten Themen mit Interesse für die Zielgruppen (Elektronische Dossiers); Ansatzpunkte dafür z. B.: gespeicherte Benutzerdaten, aktueller Bezug der Themen, Anfragen an Experten, Themen der Diskussionsrunden
- ◆ Zusammenstellung von Informationen bzw. im WWW vorhandener Informationsressourcen anhand der Ergebnisse der Informationsbedarfsanalyse

Die Interaktions- und Kommunikationsdienste des Marktplatzes sollen gewährleisten, daß der potentielle Nutzer nicht nur Informationen "abrufen" kann, sondern durch Kommunikation und Interaktion zusätzlich Möglichkeiten bekommt, Informationsprobleme zu lösen und Einfluß auf die Gestaltung des Angebotes zu nehmen. Vorgesehen werden sollten zu diesem Zweck:

- ◆ Mitgestaltungs- und Beteiligungsmöglichkeiten durch das Übermitteln von Links, Terminen usw. mittels öffentlicher Formulare
- ◆ Kommunikationsangebote (thematisch orientiert, moderiert)
 - ◆ Diskussionsforen
 - ◆ Chats
 - ◆ Gästebuch
 - ◆ Feedback / Mail-To-Funktion
 - ◆ FAQs

zusätzlich: Etablierung von geschlossenen Benutzergruppen durch die Unterscheidung von "Public-" und "Private-Bereichen"

Als *Value Added Services* können angeboten werden:

- ◆ Kontaktangebote zu Experten bei bestimmten Fragestellungen in Form einer thematisch strukturierten Übersicht über Experten und / oder als Expertenbörse, in die sich jedes Mitglied eintragen kann
- ◆ Auf die Benutzung des Marktplatzes bezogene Kontaktangebote: Was finde ich auf dem Marktplatz? Wie navigiere ich auf dem Marktplatz? Welche Quellen sind interessant?
- ◆ Individuelle Informationsanfragen zu konkreten Informationsproblemen
- ◆ Personal Communications: z. B. werden ausgewählte Beiträge aus individuellen Informationsanfragen als "neues Wissen" den anderen Marktplatzbesuchern zur Verfügung gestellt
- ◆ Push-Konzept für neueste Nachrichten aus bestimmten Bereichen.
Langfristig: Filterung des gesamten Informationsangebotes.

Die Transaktionsdienste des Marktplatzes Bildung sollen letztlich dem Tatbestand gerecht werden, daß allein die Information über vorhandene Produkte und weiterführende Informationsquellen nicht ausreichen wird, die Leistungsanforderungen potentieller Nutzer zu befriedigen. Transaktionsmodi, die der Marktplatz Bildung verwirklichen sollte, sind:

- ◆ die Recherche in eigenen (und fremden) Datenbanken bzw. Zugriff auf eigene Datenbestände; auch durch Vermittlung von Recherchen bei anderen Hosts und Übernahme der Abrechnung

- ◆ der Abruf von Volltexten bzw. auf Grundlage eines Subskriptionsmodells, Abruf elektronischer Fachzeitschriften
- ◆ Dokumentenlieferung / -bestellung: Email / Fax-Übermittlung von Volltexten
- ◆ die Übermittlung von Recherche- oder Beratungsaufträgen
- ◆ Bestellmöglichkeiten (bei anderen Anbietern) für mit dem Informationsangebot in Verbindung stehenden Produkte

Aufgrund des entworfenen Produkt- und Dienstleistungsportfolios sollte ein erster Prototyp die Beurteilung der Attraktivität und Funktionalität des Elektronischen Marktplatzes für Bildung verbessern und eine erste Evaluation ermöglichen. Abbildung 5 zeigt die Einstiegsseite des Dienstes. Deutlich wird hierbei vor allem die Struktur des Dienstes, die die genannten Informations-, Kommunikations- und Transaktions- / Interaktionsdienste auf verschiedenen Ebenen so anordnet, daß die spezifischen Nutzenerwartungen der potentiellen Zielgruppen berücksichtigt werden. Auf der Homepage des Marktplatzes sollte unmittelbar verdeutlicht werden, welche Informationen und Dienstleistungen im Angebot zu finden sind und welche Möglichkeiten sich für den Besucher bieten, das Informationsangebot zu erschließen.

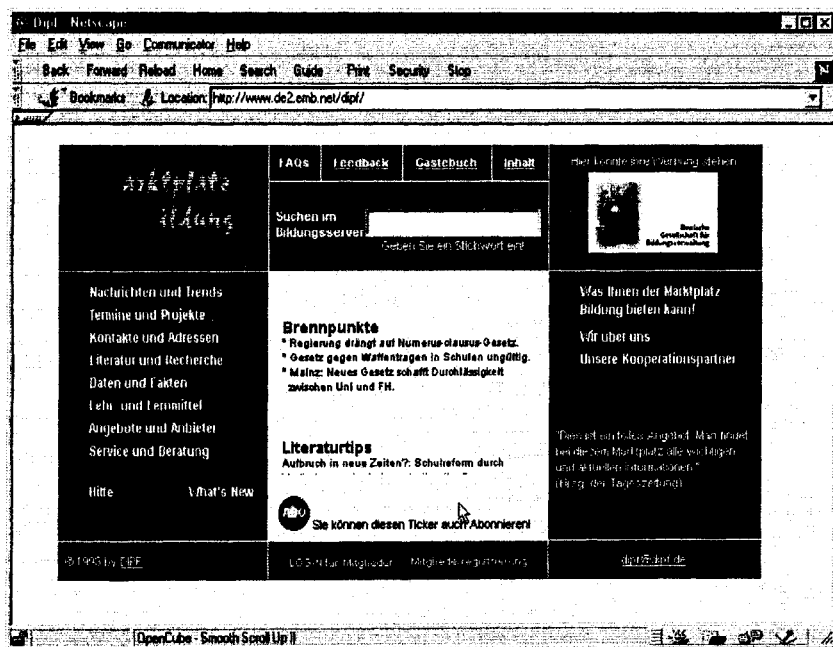


Abb. 5: Homepage des Elektronischen Marktplatzes Bildung

10 ZUSAMMENFASSUNG UND AUSBLICK

Die beschriebene Studie ergab deutliche Hinweise darauf, daß der Bedarf der potentiellen Nutzer nach einem elektronischen Fachmarktplatz, der alle für die Erfüllung fachlicher Aufgabenstellungen erforderlichen Informationen an einer Stelle zusammenführt, vorhanden ist. Sie ergab aber auch Hinweise für den Betreiber eines elektronischen Marktplatzes Bildung, welche Inhalte und welche Dienstleistungen angeboten werden können.

Dabei zeigt sich für das DIPF, daß mit Hilfe eigener Kompetenzen eine Reihe von Dienstleistungen für den Elektronischen Marktplatz Bildung erarbeitet und angeboten werden können. Insbesondere der beachtliche Wissenstand im DIPF in den Bereichen Informationsmethodik und Literaturinformation sollte zum Nutzen des Elektronischen Marktplatzes Bildung eingesetzt werden. Es bleibt aber trotzdem der Bedarf an neu zu schaffenden Informationsprodukten, welche die Attraktivität des Elektronischen Marktplatzes Bildung erhöhen, bestehen. Mit diesen soll dem Nutzer eine kompetente Anlaufstelle für fachliche Informationsprobleme präsentiert werden. Gefragt ist vor allem Fakten- und Übersichtsinformation und die verdichtete Aufbereitung von relevanter Information und wissenschaftlichen Erkenntnissen. Dieses Überblickswissen muß bei Bedarf auf dem Elektronischen Marktplatz vertieft werden, sei es durch das Angebot von Volltexten oder mit einer Beratung durch einen Mitarbeiter des Elektronischen Marktplatzes Bildung.

Literatur

[Herzog 1997]

R. Herzog: Rede von Bundespräsident Roman Herzog auf dem Berliner Bildungsforum am 5. November 1997 im Schauspielhaus am Gendarmenmarkt.
URL (4. 8. 1998):
<http://www.bundespraesident.de/reden/deutsch1997.htm>, 1997.

[Kuhlen 1996]

R. Kuhlen: *Zur Virtualisierung von Regionen durch elektronische Marktplätze*. Konstanz Informationswissenschaft, Universität Konstanz, Bericht 79-96, 1996.

[Kuhlen / Zhang 1997]

R. Kuhlen / Z. Zhang: *Building web-based scholarly communication forums using electronic journals*. International Symposium on Research, Development & Practise in Digital Libraries: ISDL '97, Tsukuba Science City, Japan, November 18-21, 1997, 1997.

[Kunz et al. 1997]

W. Kunz / H. Rittel / W. Schwuchow: *Methods of analysis and evaluation of information needs. A critical review*. Verlag Dokumentation: München, 1977.

[Schmid / Lindemann 1997]

B. Schmid / M. Lindemann: *Elemente eines Referenzmodells Elektronischer Märkte*. Sonderdruck anlässlich der WI' 97. Tutorium "Elektronische Märkte", 1997 (Bericht IM HSG/CEM/44).

Between the lines: Newsgroups as an information source for social historians

Sylva Simsova

Mrs Sylva Simsova
DATA HELP, 18 Muswell Avenue, London N10 2EG
simsova@simsova.demon.co.uk

Contents

- 1 Introduction
- 2 Use of written material as a source of data
- 3 Mass-Observation methods
- 4 Oral history methods
- 5 Analysis of texts by the content analysis method
- 6 Analysis of texts by the qualitative research method
- 7 Methods chosen for this research
- 8 The sample
- 9 The classification of articles
- 10 UK material
- 11 Contributors
- 12 Contributions
- 13 Content analysis
- 14 Conclusion

Summary

The aim of this pilot study was to show the potential usefulness of Usenet discussions to future social historians and to explore possible methods for analysing the data. A survey of 42 Usenet newsgroups was made in November 1996 to explore the patterns of information formed by the contributors. The 5541 articles written by 2130 persons were analysed to identify UK participants and topics related to the UK; to measure the frequency of contributions and relate it to their authors' interests; to draw a comparison between moderated and non-moderated groups; and to indicate the most commonly discussed topics, using keywords taken from headers. A content analysis was used to check adherence to a point of netiquette and to find out to what extent two

selected themes (the forthcoming election and violence in society) occupied the minds of contributors.

1 Introduction

Contributions to Usenet discussion groups will, if we succeed to reliably archive them for posterity, represent a primary source of information to future social historians in the same way as written correspondence and diaries.

This article presents a case study of what selected Usenet groups reveal about the communication behaviour of their participants and about social life in the United Kingdom in November 1996.

2 Use of written material as a source of data

Personal diaries and correspondence, including contributions to the Usenet discussion groups, come under the sociological category of "uncontrolled writings" (Vernon 1939) as distinct from, for instance, replies to questionnaires. Their free text makes it difficult to quantify their contents.

Further, the text does not aim to be objective. As one of the founders of Mass-Observation put it, "they tell us not what society is like, but what it looks like to them". (Madge 1937-38)

Chance plays a part in the availability of material to the researcher, for instance some letters get lost and only some find their way into an archive.

The authors of a collection of documents do not constitute a sample, strictly defined according to a scientific method, they are either self-selected (contributions to the Usenet) or selected by accidents of history (correspondence in an archive).

3 Mass-Observation methods

Mass-Observation was started in response to a letter in the *New Statesman* in 1936 which suggested that "there was considerable scope for an anthropological study of our own civilisation". (Richards 1987)

Its interdisciplinary approach was well suited to the time when "the boundaries of the key disciplines of sociology, anthropology and economics were still sufficiently fluid." (Stanley 1990)

The main method used by Mass-Observation was a "juxtaposition of many observed details of everyday life to build up a composite picture". (Mercer 1989)

The amateur observers used by Mass-Observation were members of a panel of volunteers recruited from respondents to advertisements and from informal contact. It was not claimed that it was quantitatively representative of the British population, but it did cover a wide range of age, occupation and opinion. (Mercer 1989)

Its methods were criticised by mainstream sociology for its quantitative shortcomings and sampling. (Mercer 1989)

4 Oral history methods

While Mass-Observation was interested in people's attitudes to current events, oral history was started a couple of decades later as a tool for studying the past. It arose from an interest in local history. Testimonies of witnesses were recorded to document past events which would otherwise have gone unrecorded.

At the collecting stage oral history is not written down. It is recorded on tape and transcribed afterwards.

Oral history methods aim to record facts rather than their meaning for those experiencing them. It studies the whole text rather than individual words.

5 Analysis of texts by the content analysis method

Content analysis was started at about the same time to identify the often hidden meaning of texts. It originated in media studies. (Berelson 1952, Stone 1966)

The samples for analysis consist of written texts. The actual content is considered less important than the words used to express it.

Researchers establish a set of categories and decide on words which belong to the categories. The words are tagged and their frequency counted. (Silverman 1993)

This quantitative method can extract subtle meanings which would otherwise escape researchers studying merely the texts' subject content.

6 Analysis of texts by the qualitative research method

Where historians study the content of recorded messages, researchers using qualitative methods "seek to identify the meaning of the message as a whole.."(Keele 1995)

One of the methods resembles that of content analysis. It consists of breaking up the text into sections, perhaps even words which represent the participants' categories.

The method differs from that of content analysis in using categories derived from the texts rather than those predefined by the researcher.(Coffey 1996, Silverman 1993)

7 Methods chosen for this research

The methods chosen for this research resemble those of the Mass-Observation and of content analysis.

The framework of the sample was defined by the researcher (newsgroups, number of days), but the contributors were self-selected (those who happened to write to the newsgroup on the particular day).

The Mass-Observation gathering of data was restarted in the 1980s. A panel of new volunteer contributors was recruited through the national and local press. (Sheridan 1991)

It is possible that the method of downloading Usenet contributions suggested in this exploratory study could be used to supplement the data collected by letters, questionnaires and diaries in the traditional way.

8 The sample

About 200 contributions or a week's output from each of the 42 selected groups were downloaded. A full list can be found in Table 4.

The groups were selected from over 2000 available on the Demon server.

From among the "uk." groups 16 were considered for inclusion. Of these 7 were rejected before downloading (uk.answers, uk.events, uk.finance, uk.media, uk.org.community, uk.community.social-housing) and 3 after (uk.people.deaf, uk.people.disability, uk.d-i-y). The reasons for rejection were low numbers of contributions, excessive specialisation or duplication of other material.

In the "soc.culture." groups 7 were considered for inclusion and 1 rejected after downloading (soc.culture.multicultural). The reason for rejection was that there happened to be no contributions at the time.

In the "soc.history." groups 7 were considered for inclusion and 3 rejected before downloading (soc.history.war.misc, soc.history.medieval, soc.history.science). The reason for rejection was that they were considered too specialised.

In the "soc." groups 7 were considered and 1 excluded after downloading (so.rights.human). The reason was that it did not deal much with the UK. One group was retained in spite of its strong US bias (soc.veterans).

It the "alt." groups 28 were considered for inclusion, 4 rejected before downloading (alt.chrisnet.second-coming.real-soon-now, alt.war, alt.society.deaf, alt.support.disabled.caregivers), and five after downloading (alt.macdonalds, alt.politics.correct, alt.politics.greens, alt.feminism, alt.politics.nationalism). Some of the rejected groups had too few or too many entries or did not refer to the UK situation.

In the "talks." groups only 1 was considered for inclusion. The "talks." groups often have rather ephemeral contributions. The "talk.euthanasia" was selected because no other group dealt with the subject.

Finally the sample looked as follows:

The first part of the sample consisted of 6 groups from the "uk." range, 6 groups from the "soc.culture." range, 4 groups from the "soc.history." range, and 6 groups from the "soc." range dealing with the issues of the day. All of these groups were selected to represent mainstream thinking because they were moderated.

The second part of the sample consisted of 20 groups from the "alt." range. Not being moderated they represented alternative thinking.

The number of articles in the groups varied between 3 and 289, the average being 133. The total number was over five and half thousand.

The sizes of the files containing the downloaded newsgroups were between 11k and 1360k, with an average of 389k.

The average length of articles was between 1k and 6k. In one group the average was 44k due to the inclusion of lengthy reports.

9 The classification of articles

The downloaded articles can be classified in two ways: by the level of editorial control and by subject matter.

The "uk.***" and "soc.***" newsgroups were subject to editorial control, the "alt.***" newsgroups were not edited. The unedited ones contained errors and ephemeral material which would not be so useful to historians. On the other hand they revealed spontaneous uninhibited thinking on the part of the contributors.

The subject matter fell roughly into six categories: 8 dealing with culture (2 in "alt.***", 6 in "soc.***"), 9 with environment (8 in "alt.***", 1 in "uk.***"), 5 with feminism (2 in "alt.***", 3 in "soc.***"), 7 with history (3 in "alt.***", 4 in "soc.***"), 3 with old people (2 in "soc.***", 1 in "talk.***"), 10 with politics (4 in "alt.***", 1 in "soc.***", 5 in "uk.***").

10 UK material

Identification of material dealing with the UK was not easy as the English language newsgroups were used by people from all English speaking countries. Three indicators of UK material were identified:

The first indicator was the newsgroups with "uk.***" in the title. By implication material accepted by them should deal with the UK.

The second indicator was the email addresses of correspondents. Increasingly, however, internet users in the UK are using international hosts and cannot therefore be identified from their addresses.

About 21% of email addresses had a ".uk" ending. The majority (76%) were in the "uk.***" newsgroups, a small number of them (16%) in the "soc.***" newsgroups and hardly any (6%) in the "alt.***" newsgroups.

Political newsgroups had the largest number of ".uk" email addresses (44%), followed by the newsgroups dealing with culture (24%).

The third indicator was the frequency of names of parts of the United Kingdom used as words in the contributions: UK(187), Britain (200), England(234), Scotland(188), Wales(133), Ulster(13). About half of the instances were found in the "soc.***" newsgroups and only 37% in the "uk.***" newsgroups, the reason presumably being that it was taken for granted that discussion of any subject in the "uk.***" newsgroups referred to the UK situation.

The names of parts of the United Kingdom were used most frequently in the cultural newsgroups (42%).

11 Contributors

The total number of contributors was 2130, of whom 1225 contributed one article, 787 between 2 and 7 articles (i.e. up to 1 a day), 80 between 8 and 14 articles (i.e. up to 2 a day) and 38 between 15 and 55 articles (i.e. up to 3 or more a day).

The single article contributions were mostly in the "soc.***" newsgroup and the "alt.***" group. The most frequent contributors sent their articles to several

different groups. The most frequent combination of subjects was culture and politics.

One could roughly say that the "uk.***" groups were written to at a more or less constant frequency, while the frequency of "alt.***" and "soc.***" decreased with the number of postings and the mixed category increased with the number of postings.

There were 21 contributors who wrote on average between 3 and 8 times a day. These were investigated in depth.

They posted most frequently to UPMIS (19% of the 714 articles) and APEQA (14%). They also posted to UPELE and ASEAR.

From the address and the organisation name accompanying their articles they could be classified into three types : academic, computer and communication staff, private. The highest number of contributors was in the private category (43%) and among the computer and communication staff (38%). Only 19% were from the academic world.

According to their address 48% were from the UK.

The enthusiasm of some of these correspondents could be indicative of a permanent feature of their life, though some of them wrote only occasionally during temporary leisure. The email addresses of the top 11 correspondents were checked in September 1997 to see if they were still writing as ardently as a year before. Bearing in mind that the comparison could be influenced by the dejanws not indexing all the newsgroups, the result for one week was as follows: 3 contributors did not write anything, 3 wrote between 2 and 4 times , 2 between 10 and 13. Only three maintained their high rate of postings between 23 and 28 articles a week.

12 Contributions

The sample contained 5541 articles. One quarter dealt with politics, another quarter with culture, yet another quarter with social problems, such as women or the elderly, one eighth with history and one eighth with the environment.

The articles had 1836 headers of which 1502 were used in one group, 126 in two groups, 20 in three groups, 4 in four groups and 1 in five groups.

Most headers were used in newsgroups of the same category ("u.***" or "a.***"). Some correspondents posted across the boundaries between categories : 14 were in the "u.***" and "s.***" newsgroups and 14 in the "s.***" and "a.***" newsgroups. Only 1 was in the "u.***" and "a.***" groups. This indicates a commonality of interests between the participants of the various groups.

Tables 1 and 2 show the most frequently used headers for all the groups combined and for the groups with a specific UK interest. The score was calculated as a percentage of all headers in the group or combination of groups. Table 3 gives the terms used in the headers. They are arranged in rough subject groups and provide an indication of the range of contributors' interests.

A brief listing of the top two or three under each subject group, extracted from the table, suggests that the minds of contributors were filled with forebodings and unhappy memories:

culture: race, slavery
environment: barefoot, vegetarians
feminism: abortion, marriage
history: holocaust, Halloween, Irish holocaust
medicine: suicide, aids
old: retirement, euthanasia
politics: election, IRA, anarchy
psychology: violence, responsibility
recreation: tartan, food, knotwork
religion: Christian, Fall
science: climate prediction, toxicity
society: victim disarmament, gun, arrest

13 Content analysis

Using the content analysis method, the text of the articles was subjected to a word count of selected words to ascertain three points:

First, to what extent did the contributors abide by the rules of netiquette. Two swearwords were selected ("s" and "f") and their occurrence in the sample counted:

About 9% of articles in the "s.*" newsgroup and the same percentage in the "a.*" newsgroups contained one of the two words. Analysed by subject matter the feminist newsgroups had the highest incidence of the words (21% of articles), followed by the old people newsgroups (13%) and politics (10%). The highest incidence was in SMEN (73%), followed by AANGE (54%) and ANGS (41%).

Second, to what extent did the contributors show an interest in the then forthcoming election. The occurrence of the word "election" in the sample was counted.

About 49% of articles in the "u.*" newsgroups used the word "election". In the "a.*" newsgroups it was used by only 12% of the articles and in the "s.*" groups by only 4%. Analysed by subject matter the highest incidence of the word was in politics (35%), followed by environment (22%). In the feminist newsgroups the incidence was very low - 1% only.

Third, a group of words was selected to investigate the theme of violence: gun, kill, murder, bomb. As the discussion on the banning of guns was in progress at the time the incidence of these words was rather high.

In the "u.*" newsgroups the words were used by three quarters (74%) of the articles. In the "a.*" and "s.*" groups they were used by 20% articles in each. Newsgroups dealing with politics had the highest incidence of the violence words (54%), followed by history (29%) and feminism (28%). Two newsgroups had more than one word per article : UPCON (165%), SHLIV (110%). Other newsgroups which scored high on the theme of violence were SMEN (87%), UPMIS(77%), UENV(73%). The newsgroups which scored low on the theme of violence were UPELE (2%), SCWEL (3%), AWSUP (3%), ALBAR (3%). Three newsgroups scored zero: SCCOR, SPOL, SRET.

Due to the lack of equipment, the methods of content analysis could only be used in a simplified version. For this exploratory research the count was done by the search/replace command of the word processor.

14 Conclusion

The purpose of this survey was threefold:

Firstly, to draw attention to a seemingly ephemeral method of communication and to its value in the stream of the collective memory. With fewer people writing personal letters or "letters to the Times", the Usenet may eventually become a rare source of information about what people in our time talked about.

Secondly, to provide some data about the behaviour of contributors on the Usenet. The frequency of postings and cross-postings suggests an information net created spontaneously by the users. There is a need for further case studies to establish the most common patterns of information behaviour.

Thirdly, to show how the method of analysing Usenet contributions relates to similar methods used elsewhere, particularly in Mass-Observation and in content analysis. It is hoped that this pilot study will point the way for further studies of the Usenet on a larger scale and with better technical equipment. (Simsova 1995, Simsova 1996)

References:

- [Berelson 1952]
BERELSON B.R. *Content analysis in communications research*. The Free Press, 1952.
- [Coffey 1996]
COFFEY A. and ATKINSON P. *Making sense of qualitative data...* Sage, 1996, p.5.
- [Keele 1995]
KELLE U. ed. *Computer-aided qualitative data analysis*. Sage, 1995 p.167.
- [Madge 1937-38]
MADGE C. and HARRISON F. *First year's work 1937-38 by Mass Observation*.
Lidsey, Drummond, 1938, p.66.
- [Mercer 1989]
MERCER N. *Mass-Observation 1937-40...* Univ. of Manchester Faculty of Economic and Social Studies, 1989, p.3,9,56.
- [Richards 1987]
RICHARDS J. and SHERIDAN D. *Mass-observation at the movies*. Routledge, 1987, p.2
- [Sheridan 1991]
SHERIDAN D. *The Tom Harrison Mass-Observation archive...* Univ. of Sussex Library,
1991, p.3-9.
- [Silverman 1993]
SILVERMAN D. *Interpreting qualitative data...*Sage, 1993, p.10.
- [Simsova 1995]
SIMSOVA S. *Representation of cultures on the Usenet : a survey*.
Online Information 95 Proceedings, p 437-454;
- [Simsova 1996]
SIMSOVA S. *Some characteristics of the users of Czech mailing lists dealing with librarianship and computer science*. Focus on Internat. and Comparative Librarianship vol 27, no 3, Dec 1996, p 154-158.
- [Stanley 1990]
STANLEY L. *The archeology of a 1930s Mass-Observation project*. Univ. of Manchester Dept. of Sociology, 1990, p.37.
- [Stone 1966]
STONE P.J. *The General enquirer: a computer approach to content analysis*. Cambridge, Ma.: The MIT Press, 1966
- [Vernon 1939]
VERNON P.E. in BARTLETT, F.C. *The study of society*. 1939, p.199-200.

Table 1 Top headers for all newsgroups combined - 20% and over

newsg	tot	items	
1	223	20.5	45 Not Same Species Marriages???
1	78	20.5	16 vegetarians eating meat????
1	48	20.8	10 Right to Die
1	48	22.9	11 Assisting suicide is malpractice
1	21	23.8	5 Where in time should I start?
1	78	24.4	19 Engineered Soya? [was Bombay duck]
2	269	24.9	67 An Englishman's home is no longer his castle
2	88	25.0	22 UK proposed ban on handguns -reactions?
1	262	25.6	67 Gun Control after Dunblane
1	46	26.1	12 Interested in Reenactment Groups
1	25	32.0	8 Land For Building In Cornwall
1	218	44.0	96 Victim disarmament in the UK
1	73	45.2	33 Skull in Boulder images
1	40	50.0	20 Where did you *barefoot* today?
2	45	57.8	26 After the Fall
1	204	69.6	142 Thoughts on Abortion

Table 2 Top headers for "uk.*" and "soc.*" - 10% and over

newsg	tot	items	
UPELE	207	11.1	23 If it is 1st May.....
SCCEL	204	11.3	23 Celtic Knotwork prior to Christianity
SCCOR	25	12.0	3 Hal an Tow
UENW	78	12.8	10 Foxhunting
SCSCO	220	13.2	29 Why Tartan and Other Questions
UGLOC	15	13.3	2 Spontaneous Marxist demos in Hamburg
SCBRI	206	13.6	28 Violent America [UK Bans handguns]
UPCON	218	13.8	30 UK SHOOTERS SHAFTED BY CHARLATANS!
UPCON	218	15.6	34 UK Pistol Ban for America Awaits Tragedy
UPELE	207	19.8	41 Election Law Reform?
UENW	78	20.5	16 vegetarians eating meat????
UENW	78	24.4	19 Engineered Soya? [was Bombay duck]
UPMIS	262	25.6	67 Gun Control after Dunblane
SCCOR	25	32.0	8 A Question About Land For Building In Cornwall
UPCON	218	44.0	96 VICTIM DISARMAMENT IN THE UK
UPENV	63	84.1	53 An Englishman's home is no longer his castle

Table 3 Terms used in headers - part1

cul	98		fem	839		god	182
cul	49	race	fem	146	abortion	god	39 Christian
cul	28	slavery	fem	100	Marriage	god	28 fall
cul	5	Jews	fem	91	women	god	12 flood
cul	4	skin colour	fem	86	wives	god	10 atheist
cul	3	stereotypes	fem	53	feminism	god	9 God
cul	3	nationalism	fem	52	men	god	8 Mormon
cul	2	language	fem	46	genital mutilatn	god	7 prophecy
cul	2	diversity	fem	43	conception	god	6 survivalist
cul	2	black	fem	29	homeless	god	6 beast
env	73		fem	28	rape	god	5 Trinity
env	26	barefoot	fem	24	sexism	god	5 satan
env	17	environment	fem	15	child	god	5 Old testament
env	16	vegetarians	fem	14	male bashing	god	5 evolution
env	4	trees	fem	10	polygamy	god	5 Bible
env	4	greenpeace	fem	9	status of women	god	4 pope
env	2	hedgerow	fem	9	sex drive	god	4 nativist
env	2	earth	fem	9	lesbian	god	4 good girls
env	2	conservation	fem	9	lovers	god	2 psalm
			fem	8	spanked	god	2 prayer
			fem	8	erotic	god	2 pagan
			fem	5	porn	god	2 Muslim
			fem	5	honoring women	god	2 Jesus
			fem	4	naked	god	2 faith bashing
			fem	4	College	god	2 evil
			fem	4	clitoris	god	2 catholicism
			fem	3	seduction	god	2 blasphemy
			fem	3	male intrigue	god	2 baptism
			fem	3	fertility		
			fem	3	domestic violence		
			fem	2	nags		
			fem	2	gay		
			fem	2	friends		
			fem	2	fashion		
			fem	2	dating		
			fem	2	dad		
			fem	2	circumcision		
			fem	2	boys		

Table 3 Terms used in headers - part 2

his 257	old 19	psy 192
his 27 holocaust	old 7 retirement	psy 39 violent
his 26 Halloween	old 6 euthanasia	psy 31 responsibility
his 24 Irish holocaust	old 4 veterans	psy 13 logic
his 22 history	old 2 pre-retirement	psy 11 headstone
his 18 druids		psy 10 humanism
his 17 Hansard	pol 379	psy 9 seniors
his 16 masonic	pol 46 election	psy 9 creationism
his 15 RU-486	pol 45 Tories	psy 9 accent
his 12 reenactment	pol 42 IRA	psy 8 Cook
his 12 Nazi	pol 28 anarchy	psy 6 insanity
his 9 civil war	pol 22 liberals	psy 6 angst
his 6 famine	pol 21 natural law	psy 5 control
his 5 ice age	pol 20 proportional repr	psy 4 survival
his 5 historian	pol 20 byelection	psy 4 self abuse
his 5 colonialism	pol 14 party	psy 4 independence
his 5 potato famine	pol 12 European Union	psy 3 tolerance
his 4 song	pol 10 war	psy 3 no touch
his 4 Roswell	pol 10 socialism	psy 3 behaviour
his 4 Gunpowder Plot	pol 10 draft	psy 3 alone
his 3 Salem Witch	pol 8 president	psy 2 tragedy
his 2 Zeus	pol 7 civil rights	psy 2 silence
his 2 mythology	pol 6 representation	psy 2 moral decision
his 2 Gulf war	pol 5 vote	psy 2 missing
his 2 enlightenment	pol 5 Labour	psy 2 exploring
his 2 clan	pol 5 equality	psy 2 death
his 2 civilization	pol 4 right-wing	
his 2 civilisation	pol 4 demos	
his 2 braveheart	pol 4 democracy	
his 2 archeology	pol 4 communist	
	pol 3 PC bigots	
med 62	pol 3 EuroCon	
med 14 suicide	pol 2 royalfamily	
med 12 medical	pol 2 policies	
med 9 aids	pol 2 peace	
med 7 hearing aid	pol 2 oppression	
med 4 placenta	pol 2 constitution	
med 3 malaria	pol 2 capitalist Party	
med 3 Health care	pol 2 alternative gov	
med 2 sick	pol 2 affirmative act	
med 2 plagues		
med 2 hunger		
med 2 hormone		
med 2 cigarette		

Table 3 Terms used in headers - part 3

rec 210	sci 174	soc 427
rec 29 tartan	sci 43 climate predict	soc 115 victim disarmnt
rec 26 food	sci 41 toxicity	soc 68 gun
rec 25 knotwork	sci 13 phone	soc 38 arrested
rec 19 soya	sci 10 Me-109E	soc 34 pistol
rec 11 rugby	sci 10 chemical agents	soc 24 lifers
rec 10 foxhunting	sci 7 oil production	soc 16 kilt
rec 9 libraries	sci 6 reptiles	soc 14 tax
rec 9 film	sci 6 airships	soc 11 bomb
rec 7 slam	sci 5 satellites	soc 11 press relations
rec 6 watchfires	sci 5 fossil	soc 9 assasination
rec 6 olympics	sci 4 patent	soc 8 land
rec 5 bonfires	sci 4 computer	soc 8 building
rec 4 pipe band	sci 3 Toyota	soc 7 underlord
rec 4 lanterns	sci 3 ship	soc 6 marxist
rec 3 sugar	sci 2 TW800	soc 6 legislation
rec 3 slang	sci 2 science	soc 6 knives
rec 3 New Year	sci 2 rockets	soc 5 loan sharks
rec 3 motorist	sci 2 netscape	soc 5 insurance
rec 3 latin	sci 2 kitten	soc 5 finance
rec 3 holidays	sci 2 helicopter	soc 4 third world
rec 3 concert	sci 2 constellations	soc 4 police killing
rec 3 aesthetics		soc 4 justice
rec 2 wrestlers		soc 3 names
rec 2 winter		soc 3 horn
rec 2 teacher		soc 3 cost of living
rec 2 recipes		soc 3 Communities
rec 2 hurling		soc 2 terrorism
rec 2 hunt		soc 2 opinion poll
rec 2 golf		soc 2 Media bias
rec 2 fishing		soc 2 lawsuit
		soc 2 impeachment
		soc 2 generation X

Table 4 List of newsgroups in the sample

AACT	alt.activism
AANA	alt.anarchism
AANGE	alt.anger
AANGS	alt.angst
ABPRO	alt.bible.prophecy
ACAT	alt.catastrophy
ADISC	alt.discrimination
AEAR	alt.earth_summit
AFIND	alt.feminism.individualism
AHFUT	alt.history.future
AHIF	alt.history.what-if
AHLIV	alt.history.living
ALBAR	alt.lifestyle.barefoot
AOEAR	alt.org.earth-first
APEQA	alt.politics.equality
ASEAR	alt.save.the.earth
ASFUT	alt.society.futures
ASUR	alt.survival
AWSUP	alt.women.supremacy
SCBRI	soc.culture.british
SCCEL	soc.culture.celtic
SCCOR	soc.culture.cornish
SCIRI	soc.culture.irish
SCSCO	soc.culture.scottish
SCWEL	soc.culture.welsh
SFEM	soc.feminism
SHIF	soc.history.what-if
SHIS	soc.history.
SHLIV	soc.history.living
SHMOD	soc.history.moderated
SMEN	soc.men
SPOL	soc.politics
SRET	soc.retirement
SVET	soc.veterans
SWOM	soc.women
TEUT	talk.euthanasia
UENV	uk.environment
UGLOC	uk.gov.local
UPCON	uk.politics.constitution
UPELE	uk.politics.electoral
UPENV	uk.politics.environment
UPMIS	uk.politics.misc

Projekt "Infopoint"- Errichtung eines Informationskiosks an der Universität Graz

Gunter Bauer, Gerhard Reichmann

Karl-Franzens-Universität Graz
Institut für Informationswissenschaft
Universitätsstraße 15 / F 3
A-8010 Graz
email:
gunter.bauer@kfunigraz.ac.at
gerhard.reichmann@kfunigraz.ac.at

Inhalt

- 1 Einleitung
- 2 Aufstellungsort
- 3 Informationsangebot
- 4 Auswahl der Hard- und Software
- 5 Realisierung
- 6 Benutzerstudien und Weiterentwicklung

Zusammenfassung

Der vorliegende Artikel beschreibt ein am Institut für Informationswissenschaft der Karl-Franzens-Universität Graz geplantes und größtenteils realisiertes Projekt, das die Errichtung und den Betrieb eines Informationskiosks für Studierende zum Inhalt hat.

Abstract

This paper deals with a project started by the Department of Information Science at the University of Graz concerning the development and operation of an information corner.

1 Einleitung

Im Bereich (Gebäudekomplex) des RESOWI-Zentrums der Karl-Franzens-Universität Graz, der Ausbildungsstätte für ca. 15.000 Studierende der Rechtswissenschaften sowie der Sozial- und Wirtschaftswissenschaften, soll an zentraler Stelle ein sogenannter "Infopoint" errichtet werden, an dem Studierenden, Universitätsbediensteten und sonstigen Interessenten universitätsrelevante und andere Informationen angeboten werden. Zur Verwirklichung dieses Zieles wurde am Institut für Informationswissenschaft im Herbst 1997 ein Projekt gestartet, das sich in die fünf Phasen "Festlegung des Aufstellungsortes", "Bestimmung des vom Zielpublikum gewünschten Informationsangebotes", "Auswahl der geeigneten Hard- und Software", "Realisierung" sowie "Benutzerstudien und Weiterentwicklung" untergliedert. Während den Institutsmitarbeitern die Planung und Koordination der einzelnen Projektaktivitäten oblag bzw. obliegt, erfolgt(e) die Umsetzung der geplanten Schritte unter Mitwirkung von Studierenden im Rahmen von Lehrveranstaltungen des Instituts. Die folgende Abbildung gibt einen Überblick über den zeitlichen Ablauf des Projekts.¹

<u>Projektphase</u>	<u>Beginn</u>	<u>Ende</u>
Aufstellungsort	<i>Oktober 1997</i>	<i>Jänner 1998</i>
Informationsangebot	<i>Oktober 1997</i>	<i>Juni 1998</i>
Auswahl der Hard- und Software	<i>Jänner 1998</i>	<i>Juni 1998</i>
Realisierung	<i>Juli 1998</i>	Oktober 1998
Benutzerstudien und Weiterentwicklung	November 1998	offen

Abbildung 1: Zeitlicher Ablauf des Projekts "Infopoint"²

Vor dem Beginn der ersten Projektphase wurde zur Dokumentation der Sinnhaftigkeit des Gesamtprojekts das grundsätzliche Interesse der potentiellen Nutzer an einem Informationskiosk erhoben, wobei 96 von 107 befragten Personen mittleres bis starkes Interesse an der Errichtung einer solchen Informationseinrichtung zeigten.³ Somit kann man davon ausgehen, daß die Realisierung des Projektes "Infopoint" auf breites Interesse stößt und somit eine sinnvolle Investition darstellt.

¹ Stand: September 1998

² Bei den kursiv gedruckten Terminen in der Abbildung handelt es sich um bereits realisierte Termine.

³ Es handelte sich dabei um eine mündliche Befragung in Form einer Zufallsstichprobe, die im Gebäudekomplex des RESOWI-Zentrums durchgeführt wurde. Befragt wurden sowohl Studierende als auch Universitätsbedienstete.

2 Aufstellungsort

Im Rahmen der ersten Projektphase war der optimale Standort für einen Informationskiosk im Bereich des RESOWI-Zentrums der Karl-Franzens- Universität Graz zu ermitteln. Zu diesem Zwecke wurden in einem ersten Schritt im Zuge einer Gebäudebesichtigung potentielle Standorte ermittelt, wobei besonders auf die Aspekte "Umweltbedingungen", "Vandalismusgefahr", "Zugänglichkeit", "Raum" und "Technik" zu achten war. Da der geplante Informationskiosk vor allzu großer Hitze, Kälte und Feuchtigkeit bzw. vor Vandalismus geschützt sein sollte, schieden zunächst sämtliche Standortmöglichkeiten im Freien aus. Um dennoch eine möglichst gute Zugänglichkeit zu gewährleisten, wurden weiters alle Standorte in separat abgesperrten Räumen, wie etwa in Vortrags- oder Institutsräumen, aus der engeren Wahl genommen. Des weiteren wurden nur solche Standorte in Betracht gezogen, die über ausreichend Platz und entsprechende technische Ausstattung, wie z.B. Steckdosen oder einen Netzwerksanschluß, verfügten. Auf diese Art und Weise wurden zehn potentielle Standorte für den Informationskiosk gefunden, die anschließend einer "Frequenzanalyse" unterzogen wurden. Dabei ging es um die Ermittlung der Häufigkeit, mit der die einzelnen Standorte von den möglichen Nutzern aufgesucht wurden.⁴ Das Ergebnis dieser Analyse war eine Reduktion der Standortmöglichkeiten auf fünf Standorte, die auch über eine entsprechende "Frequenz" verfügten. In einem letzten Schritt wurden schließlich 265 potentielle Benutzer befragt⁵, welchen dieser fünf in Frage kommenden Standorte sie bevorzugen würden. Die Befragung führte zu einem eindeutigen Resultat: als beliebtester Standort erwies sich der Haupteingangsbereich des RESOWI-Zentrums.⁶ Für diesen Standort sprechen vor allem dessen große Bekanntheit und leichte Erreichbarkeit. Zudem ist dort eine permanente Überwachung der Benutzung möglich, da sich die Portiersloge, die während der Öffnungszeiten des Gebäudes ständig besetzt ist, ebenfalls in diesem Bereich befindet.

3 Informationsangebot

Ziel einer (teilweise) parallel durchgeführten Projektphase war es, Art und Umfang jener Informationen zu ermitteln, die am Informationskiosk angeboten werden sollten. Dazu wurde zunächst mit Hilfe eines "brainstormings", an dem sowohl Universitätsbedienstete als auch Studierende teilnahmen, eine Liste mit potentiell interessanten Informationen erstellt. Eine anschließende Befragung⁷ der möglichen Nutzer⁸ des Informationskiosks diente der Festlegung des opti-

⁴ Es wurde die Anzahl der Personen erhoben, die innerhalb eines bestimmten Zeitraumes (10 Minuten) den jeweiligen Bereich (Standort) betrat. Diese Ermittlung erfolgte mehrmals zu verschiedenen Tageszeiten. Anschließend wurde für jeden der zehn Bereiche ein Mittelwert (die "mittlere Frequenz") berechnet.

⁵ Art der Befragung: siehe Fußnote 3

⁶ Den zweiten Platz belegte dabei die im ersten Stock des Gebäudes gelegene Cafeteria.

⁷ Art der Befragung: siehe Fußnote 3

⁸ Es wurden hier 189 Personen befragt.

malen Informationsangebotes. Das Ergebnis dieser Befragung ist in Abbildung 2 in komprimierter Form dargestellt, wobei für die Bedeutung der einzelnen Bereiche Prioritätsstufen vergeben wurden: die mit 1 gekennzeichneten Bereiche sollten jedenfalls Bestandteil des Informationsangebotes sein, jene mit 2 bewerteten stoßen ebenfalls noch auf überdurchschnittlich großes Interesse; dagegen sollten die Themenkomplexe der dritten Prioritätsstufe nur im Falle eines umfangreicheren Informationsangebotes enthalten sein.

<u>Themenbereich</u>	<u>Prioritätsstufe</u>
A) Universitätsbezogene Informationen	
<u>1.) Informationen zu den Lehrveranstaltungen</u>	
- Lehrinhalte	1
- Lehrmethoden	2
- Anmeldung zu Lehrveranstaltungen	1
- Prüfungsanforderungen	1
- Literatur zu Lehrveranstaltungen	2
<u>2.) Informationen zu den Instituten</u>	
- Aktuelle Termine (Prüfungstermine)	1
- Institutsvorstellungen	2
- Prüfungsergebnisse	2
<u>3.) Sonstige Informationen</u>	
- Studienpläne / Studienführer	2
- Veranstaltungskalender (universitätsintern)	2
- Lageplan der Universitätseinrichtungen	2
- Informationen zu Auslandsstudien	2
- Geschichte der Universität	3
B) Sonstige Informationen	
- Ferial- und Nebenjobs	1
- Jobs für Universitätsabsolventen	1
- Veranstaltungskalender (universitätsextern)	2
- Wohnungsbörse	2
- Gastronomieführer	3
- Reisebörse	3

Abbildung 2: Erwünschtes Informationsangebot

Demnach werden in erster Linie universitätsbezogene Informationen gewünscht, unter denen die Informationen über Lehrveranstaltungen an erster Stelle rangieren. Von den sonstigen Informationen wird lediglich den Informationen über Jobs aller Art eine herausragende Bedeutung eingeräumt.

4 Auswahl der Hard- und Software

Im Zuge dieser Projektphase wurden für den geplanten Informationskiosk geeignete Systemlösungen ausgewählt und anschließend hinsichtlich der Eigenschaften "Benutzerfreundlichkeit", "Flexibilität", "Robustheit" und "Kosten" einem Scoringverfahren unterzogen. Allerdings konnten den Ergebnissen dieser Bewertung in der Realisierungsphase nicht Rechnung getragen werden, da die tatsächlich zum Einsatz kommende Hard- und Software von einem Sponsor zur Verfügung gestellt wird.

5. Realisierung

Das Projekt "Infopoint" befindet sich derzeit⁹ in der Phase der Realisierung. Im Laufe des Wintersemesters 1998/1999 soll der Informationskiosk jedenfalls einsatzbereit sein. In der Folge sollen verschiedene Anforderungen an einen "Infopoint" und die in diesem Zusammenhang auftretenden Probleme angeführt werden.

Wichtig für die Attraktivität eines Kiosksystems sind die Gestaltung und der Aufbau der Eingangsseiten¹⁰ (vgl. Holfelder 1995). Derartige Eingangsseiten sollten

- übersichtlich gegliedert sein (Tabellenform),
- ein auf allen Seiten gleichbleibendes Menübanner enthalten, in dem wichtige Navigationselemente ("Home", "Vor", "Zurück") zusammengefaßt sind,
- anwählbare Buttons von ausreichender Größe aufweisen,
- einen "Hilfe-Button" enthalten, über den Erklärungen zu wichtigen Funktionen abrufbar sind,
- vom Inhalt und der farblichen Gestaltung her nicht überladen sein und
- über ein einheitliches Corporate Design verfügen.

Die eigentliche Startseite könnte auch Animationen¹¹ enthalten, um einen höheren Aufmerksamkeitswert zu erzielen. Die Größe der Startseite sollte der vorhandenen Bildschirmauflösung entsprechen, damit ein Scrollen nicht nötig ist.

Bei einem öffentlichen Kiosksystem können sich im Falle einer integrierten Emailfunktion Probleme durch die Möglichkeit der anonymen Versendung von Emails ergeben. Um einen derartigen Mißbrauch zu verhindern, wird das geplante Kiosksystem über keine Emailfunktion verfügen. In einer erweiterten Ausbaustufe, bei der sich der Benutzer z.B. mittels einer Magnetkarte identifiziert, wäre eine solche Restriktion nicht mehr nötig.

⁹ September 1998

¹⁰ Beim geplanten "Infopoint" wird es sich um eine "Internetlösung" handeln, wobei einige Eingangsseiten eigens für diesen Informationskiosk erstellt werden sollen, während für die angebotenen Detailinformationen auf bestehende Internetseiten, wie etwa jene der einzelnen Universitätsinstitute, zurückgegriffen werden soll.

¹¹ Z.B. animierte GIFs, Movies von der Festplatte, Sounds.

Dafür werden Filterprogramme bereits beim geplanten Kiosksystem zum Einsatz kommen. Diese Programme überwachen den Zugriff auf das Internet. Sie arbeiten meist nach dem einfachen Prinzip, daß sie Listen von Web-Sites verwenden, die nicht angezeigt werden sollen („Negativlisten“) oder Listen von Web-Sites, die angezeigt werden sollen („Positivlisten“). Bei anderen Filterprogrammen werden die Textinhalte der jeweils zu ladenden Seiten auf das Vorkommen unerwünschter Stichwörter überprüft. Probleme mit diesen Filterprogrammen ergeben sich einerseits aus dem Umstand, daß die Listen der Web-Sites stets aktuell gehalten werden müssen (regelmäßige Updates dieser Listen sind meist kostenpflichtig) und andererseits aus der Tatsache, daß hauptsächlich nur englischsprachige Stichwörter verwendet werden. Die Überprüfung des Textinhalts einer Webseite berücksichtigt auch keinesfalls den Inhalt von Bildern, da dieser nicht klassifiziert wird. Zudem lassen sich Filterprogramme meist durch einfache Tricks überlisten (vgl. Schmidt 1997).

An dieser Stelle sollte erwähnt werden, daß es derzeit schon Bemühungen gibt, Webinhalte bereits beim Anbieter zu klassifizieren, indem Seiteninhalte mit PICS-Spezifikationen (PICS = Plattform for Internet Content Selection) versehen werden. Jedoch ist diese Klassifizierung zur Zeit noch nicht sehr weit verbreitet (vgl. Schmidt 1997, 224).

Aufgrund der Unzuverlässigkeit der angebotenen Filtersoftware¹² (vgl. Beyer 1998) soll beim geplanten Kiosksystem eine Filterung direkt beim Zugangsanbieter vorgenommen werden. Zugangsanbieter wird ein Webserver des EDV-Zentrums der Universität sein. Auf diesem Rechner ist ein Proxy Server-Programm installiert¹³, das eine derartige Filterung über einen eigenen Proxy-Port erlaubt.¹⁴

Für die Sicherheit und Zuverlässigkeit eines Kiosksystems ist stets wichtig, daß für den Benutzer ein Zugriff auf die Betriebssystemebene des Kiosks nicht möglich ist, damit er keine Systemveränderungen vornehmen kann. Das Abspeichern von Daten aus dem Web muß ebenfalls unterbunden werden; außer in ein vordefiniertes Verzeichnis (Cache-Files), das aber in regelmäßigen Abständen gelöscht wird. Außerdem sollte ein Kontrollprogramm laufen, das in regelmäßigen Intervallen den Zustand des Gesamtsystems abfragt und u.U. ein RESET durchführt und den Kiosk in einen definierten „Normalzustand“ zurückführt.

141

¹² Beispiele für Filter-Software wären etwa die Programme "Cyber Patrol", "Cybersitter", "Net Nanny", "Specs for Kids", "Surf Watch" oder auch "Parent Control".

¹³ Netscape Proxy Server 3.5: nähere Informationen dazu finden sich unter der Adresse <http://home.netscape.com/proxy/v3.5/datasheet/index.html>

¹⁴ Vorgesehen ist eine Beschränkung des Zugriffsbereiches auf den österreichischen akademischen Webbereich ".ac.at".

6. Benutzerstudien und Weiterentwicklung

Die letzte Projektphase hat die Durchführung von Untersuchungen zur Akzeptanz und Nutzung des "Infopoints" zum Inhalt. Mit Hilfe einer Internetkamera sollen stichprobenartig das Verhalten der Benutzer am Informationskiosk bzw. der Benutzerandrang beobachtet werden. Ziel dieser Beobachtungen ist es, Rückschlüsse auf die Benutzerfreundlichkeit und Robustheit des Systems bzw. auf einen etwaigen Erweiterungsbedarf hinsichtlich der Geräteanzahl ziehen zu können.

Die Häufigkeit des Zugriffs auf die einzelnen Bereiche des Informationsangebotes soll mit Hilfe einer Serverstatistik erhoben werden. Dabei ist sicherlich die Tatsache hilfreich, daß der beim "Infopoint" verwendete Proxy Server das Mitprotokollieren aller User-Transaktionen erlaubt. Mittels eines Analyse-Tools kann eine Serverstatistik generiert werden, wodurch die Anschaffung einer eigenen Analyse-Software, wie z.B. "WebChart oder "WebSuccess" (vgl. Puscher 1998), überflüssig wird. Aufgrund dieser Statistik kann dann das Informationsangebot entsprechend angepaßt werden. Von größerer Bedeutung für die Anpassung und Weiterentwicklung des Informationsangebotes sind aber sicherlich weitere Benutzerbefragungen, denn nur so können Lücken und Mängel im Informationsangebot festgestellt werden. Sofern es gelingt, einen weiteren "Infopoint" an einem anderen Standort innerhalb desselben Gebäudes einzurichten, wären auch Standortvergleiche möglich.

Literatur

(Beyer 1998)

Beyer, T.: *Gut behütet. Filter fürs Internet*, in: iX 9/1998, 142

(Holfelder 1995)

Holfelder, W.: *Multimediale Kiosksysteme*, Braunschweig/Wiesbaden : Vieweg 1995

(Puscher 1998)

Puscher, F.: *WebChart contra WebSuccess*, in: Internetworld Feb./1998, 61 - 63

(Schmidt 1997)

Schmidt, J.: *Kindersicheres Netz? Internet-Zugangsschutz für Kinder*, in: c't 15/1997, 224 - 232

Interpreting Knowledge in Information Science

Jela Steinerová

Comenius University, Faculty of Arts,
Department of Library and Information
Science, Gondova 2,
818 01 Bratislava, Slovak Republic
e-mail: steinerova@fphil.uniba.sk

Contents

1. Introduction
2. Information science at crossroads of technology and culture
3. Issues of knowledge interpretation
4. Position of interpretation in knowledge management
5. Model of information product for support of knowledge interpretation
6. New information professionals
7. Conclusion

Summary

The paper deals with new challenges of information science as social science and mutual relationships with the new discipline of knowledge management. Qualitatively new approach argues that more concentration on knowledge interpretation is needed. Knowledge interpretation links both information science as a strategical paradigm and knowledge management as corporate knowledge framework. A knowledge interpretation framework has been outlined including a series of spiral processes from perception to understanding. Based on knowledge interpretation background general model of information product has been proposed. Main components of the model (cognitive, semantic, value, communicative, technological) have been described and resulted into a proposal of a new methodology for generation of information products. It is concluded that knowledge interpretation framework and model of information product can be used for further development of theory, research and education of information professionals and knowledge managers.

1. Introduction

The paper is aimed at research of knowledge interpretation as part of knowledge management pointing to links to design and communication of

information products. The core concepts which we use in the paper are: information science, knowledge management, information society, interpretation, knowledge, information, knowledge representation.

In our view, information science is a social science merging social, human and technological issues of generation, communication and use of knowledge in society. Knowledge management can be viewed as a set of actions and processes covering identification, retrieval, creation, distribution and use of knowledge in organizations and social and subject domains. Although knowledge management is generally a part of corporate management, we would like to point to the crucial developmental links with information science, namely knowledge objectives, knowledge identification, acquisitions, knowledge storage, dissemination and application, which form the part of special knowledge culture not only within organizations, but also in wider social context of information science.

Both knowledge management and information science can be based on the role of human resources connected with knowledge, while information and communication technologies play an important and positive part in the cycle of information and knowledge transformations. Human interaction and social communication form the key to creating knowledge from information.

Information society is a society driven not only by information technologies and services, but especially by knowledge usage in social groups, subject domains and organizations. Knowledge representation can be explained as special ways of structuring knowledge aimed at its storage, retrieval and use. Cognitive research in information science shows that specific pieces of knowledge captured e.g. in various types of documents, include more meanings in various contexts. All these meanings form a basis for polyrepresentation of knowledge (Ingwersen, 1995). Since linear representation is not enough, we have to develop means for representing multiple meanings in various contexts and possible situations of knowledge use.

For the purpose of this paper knowledge can be explained as structured and generalized product of human mind including concepts, views, images, experience and know-how. Based on this information can be viewed as knowledge expressed in semiotic systems for communication with conventional meaning that can change the state of knowledge (of a recipient). Information and knowledge are then interrelated through multiple cycles of transformations.

What is then the position of interpretation within these social and human processes? I suppose that interpretation is connected with generation of meaning for specific pieces of knowledge. This process can be directed by rules and values (and their systems), which can be even institutionalized (e.g. libraries and information systems, bibliographic or cataloguing rules). Another assumption is that interpretation as "seeking meaning" (Kuhltau, 1993) should be paid more attention to. The reason is a need for seeking new ways of

response to global societal changes and communication patterns. Changes are also required in theory, education, practice, attitudes to technologies, culture and human resources. These assumptions will be elaborated on in the following text.

2. Information science at crossroads of technology and culture

New views of information science have been arisen by the concept of information society. Apart from quantitative growth of knowledge, the development of successively digitalized knowledge universe has been marked by gradual integration at both quantitative (spatial, local networking into larger geographic settings) and qualitative levels (integrating various knowledge forms, e.g. in multimedia). As a result the quality of potential knowledge use represented by its possible individualization (from the viewpoints of cognitive processes, problem types, creativity, professional and social activity) has been emphasized.

Current information science explains new relationship between reality and its reflection, which has been mediated by representation of knowledge (e.g. bibliographic records). With the development of new technologies this relationship becomes non-linear, using new tools and languages of representation, like e.g. semantic networks, productions rules, frames, etc. Many new hybrid languages of knowledge representation apply the principles of fragments of know-how reminding us of human thinking. Thus, information science follows at least two issues: 1. very complex (technological) tools for knowledge representation, and 2. very simple ways of user interaction with knowledge in terms of identification, retrieval, access. Historically, this relationship has developed as rather linear one (the more complex tools for representation, the simpler ways of interaction). But still, the main issue for knowledge management and information science has been left open, namely the issue of the knowledge use and knowledge interpretation (which can be accessed or retrieved rather easily).

It should be noted that it is not easy to explain these issues, as they are highly interdisciplinary, covering intersection of psychological, sociolinguistic and epistemological approaches. But they are vital for transition to deeper meanings and knowledge understanding. Many implicit philosophical traditions can be here discovered (e.g. objectivist, subjectivist, individualistic or socially oriented, Hjørland, 1998). This work cannot analyse them, but I believe that the dominant positivistic approaches are not sufficient. My approach combines objective knowledge principle (K. Popper) with social activity and the role of cognitive processes in interpretation. I believe that meanings are both individual and social (cultural) in cycles of interpretation. Using knowledge management terminology, these processes form a spiral changing tacit knowledge to explicit knowledge and again to tacit knowledge, etc.

I believe that from present philosophical streams global, integrated theories are relevant, because they oppose the rationalistic tradition of one (ordered and stable) universe against the multiverse as space of multiple worlds and multiple potential interpretations. And this should be true for both knowledge management and information science. While information science offers philosophical and general scientific paradigm for justification of this approach, knowledge management brings its corporate and organizational framework including social, professional, subject or product domains.

Information science integrates technological advances into new culture of human mind and social activity. That is why information science and knowledge management can be metaphorically placed at the crossroads of technologies and culture, while the use of technologies enhances the access to and use of knowledge and gradually forms a new culture of knowledge society.

3. Issues of knowledge interpretation

Interpretation is a process of seeking meaning of information through various transformations and representations. Without conventional rules and values applied to knowledge it would not be possible to generate meaning for particular situation. When information scientists apply these rules and values to knowledge, they form its context. In fact, they create the added-value to the acquired knowledge. (e.g. as corporate resource). In these contexts we can apply the semiotic principle stating that a term alone does not mean anything, the meaning is conveyed by term relationships which form the context (Saussure).

Knowledge interpretation can also be seen as a product of knowledge transformations including knowledge representation tools. So the specific creativity of information professionals is knowledge production usually based on special interpretative framework, e.g. conceptual scheme, common standards and values of measurement of prospective knowledge use. Information science should offer a basis for various frameworks of this kind. The question is how to bind together various pieces of knowledge in order to form a special comprehension viewpoint.

Following the analysis of the evolution of different ways of representing and transforming knowledge I propose a general framework for understanding knowledge interpretation. This framework emphasizes the processual approach (processes of perception, assessment, inference, validation, understanding, action) and is briefly outlined in the following paragraphs.

If we follow the theory of interpretation (P. Ricoeur) we can trace at least two dimensions of a piece of knowledge, the first dimension being the meaning, and the second one the event marked by personal experience and communicative situation (e.g. dialogue). The situation includes several components like objects, states, things and facts. Knowledge interpretation is a

spiral limited by two boundaries: 1. perception, and 2. understanding. Between these two boundaries a series of processes can be distinguished. At the beginning of understanding the process of assessment occurs (possible multiple meanings), then multitude of justification or verification processes follow. These justification processes imply the comparison of several assessed meanings (polyrepresentation) to a set of criteria derived from the event connected with the given piece of knowledge. Several cycles of comparison lead to validation and confirmation of the most probable meaning. This is connected with explanation of meaning and the following understanding (which can again be followed by explanation). Explanation is the means of mediating in the process of interpretation. Final understanding is then concluded by adoption of the meaning not only with individuals, but also with social and professional domains and organizations.

The crucial processes in this interpretation framework are assessment and validation. Logical probability of the true conclusion (concerning the meaning of a piece of knowledge) forms the basis for proving the interpretation for specific situation. As an example we can mention the text analysis, at present influenced by transfer of the characteristics of small language units (phonemes, morphemes) to larger units (sentence, paragraph).

This framework has been just outlined and should be enriched by principles of knowledge segmentation and integration including the role of analysis, synthesis, inference and action as part of knowledge production. Anyway, the framework can be used for further research in information science development for information society.

4. Position of interpretation in knowledge management

Knowledge interpretation as a series of transformations represents a basic source of creative thinking. That is why it is, in my view, so valuable for (corporate) knowledge management within the integration of external and internal (human) resources. It is so because the intersection of technologies and knowledge in its cognitive and social parts takes place in organizations and form a new culture (e.g. corporate products or services).

For this difficult task general knowledge on interpretation can be beneficial. Knowledge managers should know that knowledge interpretation includes the creation of a tentative conceptual model (e.g. by analysis of information sources) and its multiple (spiral) transformation to a final conceptual model. If one strives for efficient knowledge management, then it is necessary to take into the account not only easily-formalized processes of knowledge acquisition, storage and retrieval, but also human and semiotic basis of knowledge interpretation.

The position of knowledge interpretation can be explained by its production component. I suppose that a general term of information product could be

helpful in this context. Information product is determined as any result of creative knowledge interpretation which comprises the added value to analysed information sources. However, it is important to discriminate the level of the added value as a result of creative analytical, synthetic and interpretative processes. In information science a methodological ground for these products has been developed with typical provision of context (metainformation). As examples of these products we can name study analytical reports, state-of-the-art reports, reviews, argumentation essays.

In generation of information products we can distinguish five basic phases of interpretation:

1. perception, contact with new information,
2. selection, driven by pattern recognition and validation of criteria,
3. inference, merging of categories, derivation of meanings,
4. prediction, transfer out of limitations of interpreted information,
5. action (creation of a mind product).

In this respect knowledge interpretation includes semiotic (social, domain or professional) conventions and human issues of creative personalities like thinking, feeling and acting, learning and experience. Semiotic conventions usually refer to the categories of size, spatial placement and orientation, static and dynamic objects, blank spaces, etc.

Thus, knowledge interpretation is needed more and more for corporate development and for theoretical paradigmatic shift in information science research. Its position is visible through information products. Information products represent the stress on knowledge in information society, when new organizational structures concentrate around knowledge structures. The problem is that knowledge can never be so stable that its structures will be definite, but technologies have made it possible to change the traditional library concentration on physical objects (documents) and it is the first step to deep institutional changes in communication and information processes. Traditional classification and indexing theory can help form the basis for knowledge structures studies, changing the tacit knowledge to explicit knowledge.

5. Model of information product for support of knowledge interpretation

Based on this background I have developed a model of an information product as a result of knowledge interpretation. The model is composed of three viewpoints: representation of meaning, significance of cognitive processes, social standards and values in professional and social communities.

Information product determined by the use of knowledge by users of information systems and services includes five components:

- cognitive component
- semantic component
- value component

- communicative component
- technological component.

These components form the model of information product (fig.1) in the framework of knowledge management. It is based on an assumption of the maximum level of creative interpretation in knowledge production. All the components in holistic cooperation produce the context for better use of knowledge. The cognitive component comes from cognitive processes, knowledge typology and cognitive styles including their emotional and social part. The semantic component includes especially conventions of conceptual structuring and of sign assignment and perception. These conventions are dependent on relationships between the meaning and the situation (event). The component of values is connected with possible ways of knowledge use and the degree of context. The interpretation framework described above leads to the added value as the focus of the value component. The communicative component includes communicative styles connected with the design of information product for particular professional, social or subject and domain groups. As opposed to the traditional provision of information the emphasis is laid on the support of problem solving and decision-making for special user groups. The technological component supports knowledge representation and design based on features of segmentation, ordering, use of patterns, etc. in order to make the perception and use of knowledge easier, efficient and beneficial to corporate and social development.

The technological component of the model is the most dynamic part and changes are more difficult in direction upwards to cognitive component. As for the level of knowledge interpretation, the deepest level can be found with the cognitive component and through other components this level transforms to simpler, surface forms. The model can be validated through other significant criteria like knowledge representation (e.g. simpler ways with interactive components, more complex towards cognitive components), degree of analysis and synthesis, knowledge transformations, etc.

Based on the model I have also developed a methodology for the information products generation (Steinerová, 1998). Here, the components are represented by series of questions (for information professionals as producers of information products). For example, how to define the problem or subject, to set the criteria and the gist, how to identify conceptual structure (the cognitive component), or what is the insight, validity, contradictions, quality, viewpoint, etc. (the value component), or what is the level of subject coverage and user inclusion, how to transform knowledge, what is the context, relationships (the semantic component), or what are the communicative conventions, preferred user stereotypes, how to rearrange the conceptual structure (the communicative component), or what design and style is beneficial for the given situation, how to arrange, filter, visually represent, etc. (the technological component).

I regard this model as part of a new methodological framework for information science. Evolution of information science to a new paradigm shows that the cognitive, historical and social approaches are not contradictory, but that they are part of a "multiverse" of methods for creative production of information products by new information professionals.

6. New information professionals

What are the implications of the knowledge interpretation issues for new information professionals, especially with respect to their education? First, they should be aware of the importance of interpretation for knowledge management in new information institutions. Their professionalism should be connected with particular information products which they produce in information society. I suppose that the profession is linked with creativity, with value-added processes and the use of knowledge for problem-solving in cooperative social relationships.

Based on our analyses and interpretational framework the requirements for professional and personal competencies of information professionals and knowledge managers comprise two groups of competencies: 1. analytical and interpretational competencies, 2. competencies linked with interaction between external and internal knowledge. Both groups are focused on intellectual activities connected with patterns of efficient knowledge organization and use. From the perspective of knowledge management three aspects of interpretation have to be incorporated into education : 1. representation of meaning by sign systems (*semiotic approach*), 2. mechanisms of cognitive processes (*psychological approach*), 3. social rules (conventions) in social systems (*sociological approach*). It is also necessary to include skills connected with diagnosis of information needs of users within a continuous dialogue.

Technological and communicative skills form a general background for education of information professionals. General management issues are also embodied in requirements for business skills (including e.g. understanding of financial management, fundraising, entrepreneurial skills, change management, project management). Professional skills should concentrate on knowledge organization, document design and repackaging and navigation skills (for information retrieval). Understanding knowledge and social communication are, however, the most important conditions (namely the problems of generation, development, transfer, interpretation of knowledge, motivation, knowledge audit, corporate culture, human behaviour and interaction, etc.).

Education for new information professionals (e.g. knowledge managers) should not be based on technological skills, but on methodology which can develop and support creativity of knowledge use and interpretation, sound critical judgment and a lifelong ability to learn. This has been proved by surveys at our Department (1996, 1997) aimed at directions of manpower market

development in information science and knowledge management in Slovakia. Of course, the education will have to be open, flexible and modular.

A potential image of a new information professional can be described as follows: He should be able to determine the information needs including contexts of problems. He would also need to know how to set a strategy for identification and access to information sources, how to assess and select knowledge from them. He must be interested in the best ways of knowledge organization for different problems, conceptual structuring and seeking meaning. The best way and form of information packaging for the user is also significant for him. For this he needs to know about physical and intellectual mechanisms of information storage and retrieval. And for practical activities he needs knowledge about communicative techniques, the project, time and change management.

I suppose that a future information professional will be institutionally-independent (e.g. specialization in subject domains, professional groups, etc.). His activities will also include consulting and teaching integrated into support of knowledge use. He will have to understand the ways of transforming information to knowledge, characteristics of user groups, factors which determine information needs, methods of conceptual modelling and creative problem-solving.

Interpreting knowledge is the key process within the professionalism of new information professionals (knowledge managers). Knowledge interpretation as part of creative problem solving contributes highly to the corporate and social success and further development. New courses in information science should take into account these requirements.

7. Conclusion

Knowledge interpretation has not been paid much attention in information science and knowledge management so far, because it is a complex intersection of multiple aspects. However, if information science should advance in theory, the transition from formal, descriptive methods to deep, conceptual analyses and a new coherence of recorded knowledge is inevitable. The specific professionalism of future information workers will be based on generation of the added value using analyses, syntheses and interpretation. I believe that interpretation can ensure the value of information products created by information professionals.

Transition to a deeper meaning of knowledge is made possible by various types of contexts reflected in conceptual structuring as part of knowledge interpretation. Our knowledge interpretation framework and model of information product can be used for further development of education and methodology of information science. Within our curriculum we have developed new courses proposals and a textbook (Steinerová, 1998) based on these

approaches and aimed at further improvement of the information professionals education for future.

References:

Reardon, Denis F. *Knowledge Management : the discipline for information and library science professionals*. In: Papers of the 64th IFLA General Conference 1998. Booklet 7. - Amsterdam : IFLA 1998, p. 87-92.

Essers, J., Schreinemakers, J. *Nonaka's subjectivist conception of knowledge in corporate knowledge management*. In: Knowl. Org., Vol.24 (1998), No.1, p.24-32.

Hjorland, Birger. *Information Seeking and Subject Representation. An Activity-Theoretical Approach to Information Science*. Westport, Cn., London: Greenwood Press, 1997. 213p. - New Directions in Information Management, N.34.

Kuhltau, Carol C. *Seeking Meaning : A process approach to library and information services*. Norwood, NJ: Ablex, 1993. 199p.

Beghtol, Clare. *Stories: Applications of Narrative Discourse Analysis to Issues in Information Storage and Retrieval*. In: Knowl. Org., Vol.24 (1997), No.2, p.64-71.

Endres-Niggemeyer, Brigitte. *Content Analysis - a special case of text comprehension*. In: Information. Knowledge. Evolution. Amsterdam: North Holland, 1989, p.103-112.

Hjorland, Birger. *Information Retrieval, Text Composition, and Semantics*. In: Knowl. Org., Vol. 25 (1998), N.1/2, p.16-31.

Ingwersen, Peter, Willett, Peter. *An Introduction to Algorithmic and Cognitive Approaches for Information Retrieval*. In: Libri, Vol. 45 (1995), p.160-177.



Ingwersen, Peter. *Cognitive Perspectives of Information Retrieval Interaction : Elements of a Cognitive IR Theory*. In: J. Doc., Vol. 52 (1996), No. 1, p.3-43.

Hjorland, Birger, Albrechtsen, Hanne. *Toward a New Horizon in Information Science : Domain-Analysis*. In: JASIS, Vol. 46 (1995), No.6, p.400-425.

Ricoeur, Paul. *Teória interpretácie : diskurz a prebytok významu. Interpretation theory : discourse and the surplus of meaning*. Bratislava : Archa, 1997. 135p.

Steinerová, Jela. *Teória informačného prieskumu. Theory of Information Retrieval*. Bratislava: SITK-CVTI, 1996. 262p.

Fig. 1 **INFORMATION PRODUCT MODEL**

Knowledge interpretation	Components of an information product	Change
DEEP COMPLEX 	COGNITIVE COMPONENT cognitive styles types of knowledge	SLOW 
	SEMANTIC COMPONENT conventions of meanings, signs conceptual structure	
	VALUE COMPONENT degree of metainformation, context the added value	
	COMMUNICATIVE COMPONENT communicative styles, subject domains social groups, corporate, professional groups design	
	TECHNOLOGICAL COMPONENT representation and presentation styles form, medium, type of technology	
SIMPLE		FAST

Dynamische Menüs zur Recherche und Orientierung.

Konzeption der thematischen Suche für den Informationsservice „WebSite 'Methodik“

Gerhard Knorz

Fachhochschule Darmstadt, Fachbereich IuD, Haardtring 100,
D-64295 Darmstadt, Email: knorz@iud.fh-darmstadt.de

Inhalt

1 Anwendungsfall

1.1 WebSite ‚Methodik als zentrale Informationsressource für Lehren und Lernen

1.2 Architektur und Services von WebSite ‚Methodik

2 Konzeption der thematischen Suche

2.1 Problemstellung

2.2 Lösungsmöglichkeiten

2.3 Vorgeschlagenes Konzept

2.4 Gestaltungsentwurf

3 Realisierung

3.1 Vokabular und Indexierung

3.2 Datenbank

3.3 Stand der Entwicklung

Zusammenfassung

Das Konzept einer themenorientierten Suche wird vorgestellt, das gleichzeitig das Ziel verfolgt, dem Nutzer Orientierung über die inhaltliche Ausgestaltung eines Informationsbestandes zu geben. Es wurde für WebSite 'Methodik entwickelt, einen WWW-Server für Distance Learning in dual mode, und ist übertragbar auf Dokumentensammlungen mittlerer Größe, die begriffsorientiert indexiert sind. Nutzer wählen aus einem kontrollierten Vokabular Themen aus und erhalten parallel zu den entsprechend indexierten Dokumenten alle weiteren Deskriptoren angezeigt, mit der sich die Suche vom Allgemeinen zum Speziellen fortsetzen läßt. Die Häufigkeit, mit der Deskriptoren gemeinsam indexiert sind, läßt in diesem Kontext die thematischen Schwerpunkte des

Informationsservers erkennen. Die Suchfunktionalität kann in verschiedener Weise visualisiert werden. Konzeption und Aufbau der zugrundeliegenden Datenbank ist abgeschlossen, die Integration in WebSite 'Methodik ist in Entwicklung.

1 Anwendungsfall

1.1 WebSite ,Methodik als zentrale Informationsressource für Lehren und Lernen

Hochschullehre ist aus vielfältigen Gründen in die Diskussion gekommen und in Bewegung geraten. Die Unterstützung der Lehre durch neue Medien und Informationstechnik ist nicht nur ein populäres Thema, sondern wird auch mit praktischen Konsequenzen von Lehrenden und Lernenden aufgegriffen (HIS 1998). Eine im Wintersemester 1997/98 durchgeführte Befragung bei Studierenden informationswissenschaftlicher Studiengänge im deutschsprachigen Raum hat ergeben, daß „Distance Learning“ als Variante oder Komponente eines Studiums durchaus (wenngleich mit Einschränkungen) Akzeptanz findet (Beheim et al. 1998).

Der insgesamt positiven, gleichzeitig aber vorsichtigen Einstellung der Studierenden entsprechen Konzept und Praxis der Lehre im Fach Informationsmethodik im Fachbereich Information und Dokumentation der Fachhochschule Darmstadt. Der Ansatz ist als Distance Learning im Dual Mode zu klassifizieren: Zusätzlich zur Präsenzveranstaltung wird jeder Lehrveranstaltungstermin zunächst in Form einer Planung und abschließend in Form einer Dokumentation des tatsächlichen Verlaufs im Internet „gespiegelt“. Unterlagen zu Lehrveranstaltungen werden vollständig im Netz bereitgehalten. Lehrveranstaltungsteilnehmer holen sich erforderliche Unterlagen vom Server selbst ab.

Die zentrale (technische) Komponente dieser Form der Lehre im Fach Informationsmethodik ist *WebSite ,Methodik*, dessen Wurzeln bis zum Wintersemester 1995/96 zurückgehen und die unter diesem Namen und in dieser Rolle seit Sommersemester 1996 besteht (Knoez 1998). Seit dieser Zeit dokumentiert der Server lückenlos alle im Bereich der Lehre relevanten Aktivitäten (auf der Ebene der einzelnen Termine und Inhalte) und etliches darüber hinaus. Gleichzeitig stellt er auch Kommunikationsmöglichkeiten bereit.

WebSite ,Methodik besteht mittlerweile aus annähernd 500 Dokumenten (Stand August 1998), die auf ca. 2500 Dateien in ca. 200 Verzeichnissen basieren (85 MB). Die Änderungsrate beträgt, bedingt durch die notwendigen Nacharbeitungen der Lehrveranstaltungsplanung, etwa bei 20 bis 30 Dateien/Woche. Das Volumen hat sich im Laufe des vergangenen Jahres nahezu verdoppelt.

Bei dem Aufwand, den Aufbau und Betrieb von WebSite ,Methodik für eine einzelne Lehrperson ohne weitere personelle Unterstützung erfordern, muß als Ergebnis mehr als „nur“ ein Informationspool für Studierende resultieren. WebSite ,Methodik ist gleichberechtigt auch das Instrument, in dem sich alle planenden und vorbereitenden Tätigkeiten des Lehrenden abspielen, und das die Ergebnisse dieser Prozesse dokumentiert. Vorbereitung eines Vortrags oder

eines Vorlesungstermins bedeutet zunächst Recherche nach einschlägigen Materialien auf dem Server, die anschließend strukturiert, aktualisiert, ergänzt und aufbereitet werden. Gleichzeitig ist der Informationsserver auch der Ort, an dem alle neuen Informationen zur späteren Verwendung abgelegt werden.

1.2 Architektur und Services von WebSite ,Methodik

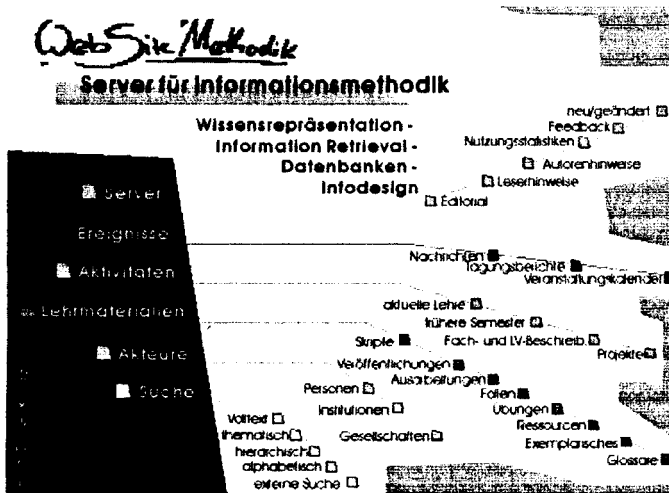


Abb. 1: Übersicht über die Rubriken von WebSite ,Methodik. Auf diese Site Map wird direkt von der Homepage aus verwiesen.

WebSite ,Methodik ist in 5 Rubriken (zuzüglich Suche) gegliedert, die weitgehend monohierarchisch, in Einzelfällen polyhierarchisch, weiterentwickelt werden (Abb. 1). Die Einordnung eines Dokumentes in die Rubrikstruktur ist zeitunabhängig und eindeutig. Die Strukturierung von WebSite ,Methodik ist bei quantitativem Wachstum und wachsenden Anforderungen stabil geblieben.

Folgende Prinzipien liegen WebSite ,Methodik zugrunde:

- **„Monotonie-Eigenschaft“ des Informationsangebotes.** Es wird grundsätzlich auf Dokumentenebene nicht gelöscht: Ein gespeicherter Link auf ein Dokument bleibt valide, da dieses Dokument entweder zeitunabhängig von Interesse ist oder aber dem Ziel dient, einen in der Zeit ablaufenden Prozeß zu dokumentieren.
- **Vollständigkeit.** Ein zentrales Qualitätskriterium einer Datensammlung ist Vollständigkeit und Abdeckung. Wann immer eine Unterlage in der Lehre verwendet oder eine Veranstaltung gehalten wird, ist sie in WebSite ,Methodik zu finden.
- **Inhaltliche und formale Qualität.** In Stichworten: inhaltliche Bewertung, verschiedener Status von Dokumenten, vorgegebene reichhaltige formale Dokumentenstruktur, Meta Tags, kontrolliertes Indexierungsvokabular.

bular, durchgängige Namenskonventionen.

- **Navigation und Suche.** Jedes Dokument ist durch seine Indexierung thematisch eingeordnet und kann über diese Indexierung auch zielgerichtet gefunden werden. Weiterhin wird von jedem Dokument aus eine Volltextsuche, ein alphabetisches und ein hierarchisches Register angeboten. Jedes Dokument listet darüber hinaus alle Standard-Pfade auf, über die dieses Dokument erreichbar ist. Ziel ist es, es dem Nutzer und seiner Zielsetzung zu überlassen, von welchen der angebotenen Navigationsmöglichkeiten er im einzelnen Gebrauch macht.

2 Konzeption der thematischen Suche

2.1 Problemstellung

Mit der Größe eines Informationsangebotes wächst der Stellenwert angemessener Navigationshilfen und Suchmöglichkeiten. In diesem Sinn hat *WebSite 'Methodik'* eine kritische Größe erreicht bzw. überschritten. Informationsangebote, die von gelegentlichen Nutzern inspiziert und punktuell genutzt werden, sind zusätzlich mit dem Problem konfrontiert, daß Nutzer erst nach und nach ein Bild davon bekommen können, zu welchen Themen Dokumente bereitgehalten werden. Außerdem können sich Themen wie „Gestaltung“ oder „Retrieval“ auf unterschiedlichen Informationsservern nicht nur quantitativ, sondern auch von der inhaltlichen Ausgestaltung her sehr unterschiedlich darstellen. Vielfach orientieren sich interessierte Nutzer zeitaufwendig per Navigation und recherchieren gezielt mittels Stichwortsuche.

Unter Nutzung der speziellen Voraussetzungen von *WebSite 'Methodik'* soll ein Konzept gefunden und vorgeschlagen werden, das eine gezielte themenorientierte Suche mit einer Orientierung über die inhaltliche Ausgestaltung des Informationsangebotes verbindet. Die Voraussetzungen betreffen

- die Größenordnung des Dokumentenbestandes (einige hundert bis einige tausend Dokumente)
- die Existenz einer Indexierung mit kontrolliertem Vokabular.

Gesucht ist also ein inhaltsorientiertes Such- und Navigationskonzept für gut erschlossene Informationsbestände, anwendbar auf klein- bis mittelgroße Dokumentsammlungen, umfangreiche Antwortmengen von Fachdatenbanken (zur Detailrecherche), oder von Subject Based Information Gateways (SBIGs), wie sie im Web in verschiedenen Projekten zwischen Browsing- und Suchdiensten positioniert werden.

2.2 Lösungsmöglichkeiten

Die Suche nach Lösungsansätzen für das gestellte Problem führt zu Arbeiten, die recht unterschiedlich motiviert sind:

Web Mining: Die Dynamik und spezifischen Anforderungen, Defizite und Möglichkeiten des WWW motivieren den Versuch, die Suche in einer mehr inhalts-

orientierten Form zu unterstützen. Besonders bekannt ist die von Alta Vista angebotene Funktionalität, dynamisch Termcluster über eine statistische Kookkurrenzanalyse in Antwortdokumentmengen zu bilden um damit für die weitere Suche Anregungen für Inclusion oder Exclusion von Wörtern und „Begriffen“ (Clustern) zu geben (Bourdoncle 1997). Weiter gehen Ansätze, in denen auch die Verweisstrukturen zwischen WWW-Dokumenten (und damit eine zwar unspezifische, aber dennoch intellektuell-basierte Entscheidung über inhaltliche Zusammenhänge) ausgewertet wird, um begriffliche Wortfelder (Stein 1997) oder Dokumentencluster (Kleinberg 1998) zu bilden.

Visual Information Retrieval Interfaces: Die graphische Darstellung und Aufbereitung von Informationsräumen hat große Potentiale, bekannte Ansätze des Information Retrieval (z.B. Clusteranalyse) mit eigenständigen Beiträgen so zu verbinden, daß die visuellen Fähigkeiten des Menschen für ein effektiveres Retrieval und ein Mehr an Orientierung genutzt werden (Rorvig 1998).

Dokumenten und Begriffs-Cluster sowie Dokument-Begriffs-Netzwerke werden als Scattermaps, als topographische Karten, als Gebirge, als Landschaften oder Räume (Däßler/Palm 1998) unter Verwendung verschiedenster Metaphern eingesetzt. So benutzt beispielsweise Lyberworld ein 3D-Raummodell auf der Basis einer Massenanziehungsmetapher (Hemmje 1993): Auf einer transparenten Kugel werden Anfrageterme interaktiv plaziert und bewegt, wobei deren „Anziehungswirkung“ auf die Dokumentenwolke im Innern der Kugel deren inhaltliche Struktur erkennbar werden läßt. Der Knowledge Browser (Däßler/Otto 1996) transformiert Antwortdokumentmengen in navigierbare 3D-Klötzchen-Landschaften.

Graphische Umsetzung Boole'scher Suchanfragen: Das gegenwärtig in der praktischen Anwendung immer noch vorherrschende Retrieval-Paradigma ist die logische Verknüpfung von Suchtermen durch Boole'sche Operatoren. Unabhängig von der Kontroverse über die prinzipielle Adäquatheit dieses Ansatzes gegenüber partial match-Verfahren (eine neuere Untersuchung dazu in (Paris/Tibbo 1998)) steht dieser Ansatz in einem paradoxem Spannungsfeld:

- Nutzer schätzen die Durchschaubarkeit und Nachvollziehbarkeit des Verfahrens (Frisch/Kluck 1997, S. 47), ...
- ... obwohl sie in der Anwendung der Operatoren objektiv Schwierigkeiten haben (Frisch/Kluck 1997, S. 42) und sich durchaus Retrievalergebnisse objektiv z.T. falsch erklären.

Es sind deshalb eine Reihe von Ansätzen entwickelt worden, um Boole'sche Suchen graphisch-interaktiv durchführen zu können, wobei einige davon Gewinne an Transparenz und Kontextinformation erreichen. Dazu gehören der InfoCrystal von (Spoerri 1994) sowie eine eigenständige Weiterentwicklung davon (Bürdek et al. 1998). Auch der Zugang zu intellektuell indextierten Dokumenten unter Anwendung von Methoden der formalen Begriffsanalyse, visualisiert durch Liniendiagramme, kann so aufgefaßt werden (Kollewe et al. 1994).

Der vorgeschlagene Ansatz muß das Problem der Data Mining-Ansätze nicht

lösen, da bereits eine begriffsorientierte Indexierung vorliegt, die unmittelbar ausgewertet wird. Sie hat von ihrem gedanklichen Ausgangspunkt einen Bezug zu (Kollewe et al. 1994), läßt sich aber einfacher umsetzen. Methoden der Informationsvisualisierung lassen sich naheliegender bei der Schnittstellengestaltung mit einbringen (siehe Abschnitt 2.4).

2.3 Vorgeschlagenes Konzept

Die Idee der thematischen Suche entwickelte sich aus der Nutzung des Schlagwortregisters in WebSite Methodik. Sie bietet an, daß der Nutzer iterativ vom Allgemeinen zum Speziellen hin sucht, und daß diese Suche gleichzeitig über das informiert, was gefunden werden kann. Die Suche liefert also Überblicksinformation über die quantitative und qualitative Ausgestaltung von Themen im vorgegebenen Dokumentenbestand.

Start der Suche: Die Suche startet entweder vom aktuellen Dokument aus, indem einer der indexierten Deskriptoren ausgewählt wird oder alternativ von einer speziellen Suchseite aus, in der die Liste aller Deskriptoren (Thesaurus) zur Selektion angeboten wird.

Ergebnis der Suche: In beiden Fällen führt die Auswahl eines Deskriptors D parallel zu zwei Ergebnissen:

- Liste aller Dokumente, die mit dem Deskriptor D indexiert sind (Abb. 2b).
- Liste aller Deskriptoren, die zusammen mit D in der Indexierung von Dokumenten vorkommen, ergänzt um die Angabe der Anzahl der Kookkurrenzen.
Tatsächlich werden zwei Deskriptorenlisten erzeugt, indem Deskriptoren, die formale Aspekte beschreiben von solchen, die inhaltliche Aspekte benennen unterschieden werden (Abb. 2a).

Iteration: Der Suchprozess kann fortgeführt werden:

- Die Selektion eines Dokumententitels führt direkt zu dem entsprechenden Dokument und beendet die Iteration.
- Die Selektion eines der Deskriptoren (z.B. von D2) führt erneut zu einem Suchergebnis, bei dem D durch (D UND D2) ersetzt wird.

Interpretation: Mit etwas Pragmatismus können wir die Dokumentenliste als die Extension des gewählten Begriffs D auf dem Server ansehen: Das Thema wird repräsentiert durch die Menge aller Dokumente, in denen es behandelt wird. Die Liste der gewichteten Deskriptoren können wir als Approximation der Intension des Begriffs D auffassen: Als Intension gilt die Menge aller Merkmale, die die Dokumente der Extension gemeinsam haben (Kollewe et al. 1994). Als Merkmale interpretieren wir die inhaltlichen Kennzeichnungen, also die zugeordneten Deskriptoren. In dieser strengen Form werden außerordentlich hohe Anforderungen an die Differenziertheit, Konsistenz und Vollständigkeit der Indexierung gestellt. Praktisch handhabbarer wollen wir demnach als Intension des Themas die Menge aller Deskriptoren ansehen, die solchen Dokumenten zugeordnet sind, die das selektierte Thema behandeln. Diese Liste von Deskriptoren,

nach Häufigkeit sortiert, gibt in der Regel einen deutlich schnelleren Überblick darüber, in welcher Hinsicht das gewählte Thema ausgearbeitet ist als die Menge der Dokumententitel (Extension). Abb. 2a und 2b sollen dies am Beispiel des Themas „Gestaltung“ verdeutlichen.

Bezeichnung	Anz	Notation
Gestaltung	19	Sac
Benutzungsschnittstelle	13	Sac
Informatik	3	Sab
Informationswissenschaft	2	Saa
Hypermedia	2	Saab
Bildbearbeitung	2	Sac
Lehrinhalt	2	Sba
Didaktik	2	Sba
Informationslinguistik	1	Saaca
Klassifikationssystem	1	Saacbb
Retrieval	1	Saacbc
Evaluierung	1	Sba

Bezeichnung	Anz	Notation
Overhead-Folie	14	Fa
Beispiel	10	Fb
Publikation	2	Fa
publiziertes Dokument	2	Fa
Lehrveranstaltung	2	Fb
Definition	2	Fcbb
Hilfe	1	Fb
Statistik	1	Fb
Richtlinie	1	Fb
WebSite Methodik	1	Fcaa

Abb. 2a: Es gibt 19 Dokumente, denen „Gestaltung“ zugeordnet ist. Dieses Thema ist eindeutig in Richtung „Benutzungsschnittstelle“ ausgestaltet (13 Dokumente). Alle anderen Begleitthemen sind randständig. Was die Form betrifft, so handelt es sich in erster Linie um Overheadfolien und Beispiele (schließt sich nicht aus!). Die dritte Spalte dient (nur) der Sortierung innerhalb eines Ranges

Titel	Datei	Ext
„Media System Design“, ein neuer Studiengang an der Fachhochschule Darmstadt (Info 7, Knorz/Seeger)	infosieben1	htm
LDV-Forum, Jg. 1998, Heft 1: Rezension R. Köhler - Schiltz: Der Diaklektometrische Atlas von Südwest-Baden (G. Schiltz)	kohler2	htm
Multimedia Ausbildung an der Fachhochschule Darmstadt. (NfD, Knorz/Seeger)	nfd1	htm
Overhead-Folie: Benutzungsoberfläche - Fallbeispiel Intranet-Anwendung	boffbi1	htm
Overhead-Folie: Design ist Interface-Design	diid1	htm
Overhead-Folie: GUI - Minimalistische Gestaltung von Pictogrammen	mgvpg1	htm
Overhead-Folie: Interaktionsparadigma: Direkte Manipulation	ipdm1	htm
Overhead-Folie: Multimedia - Manipulationsmöglichkeit durch Digitali-	ddmfbm1	htm
Overhead-Folie: Optische Wahrnehmung - Farben im Kontext	owfik1	htm
Overhead-Folie: Optische Wahrnehmung - Gestaltpsychologie	owngp1	htm
Overhead-Folie: Optische Wahrnehmung - Linien und Flächen	ownluf1	htm
Overhead-Folie: Visualisierung hierarchischer Strukturen	vhs1	htm
Overhead-Folie: Visualisierung von Dokumentmengen	vvdm1	htm
Overhead-Folie: Visuelle Gestaltung - 7 Dimensionen	sddvg1	htm
Overhead-Folie: Visuelle Metapher in 3D-Welten	vmi3dw1	htm
Overhead-Folie: Visuelle Metaphern - Klassifikation	kvm1	htm
Overhead-Folie: Zeichnen planarer Graphen: Verschiedene Verfahren	zpgvv1	htm
Rezension zu: Bonsiepe, Interface. Design neu begreifen. (Nake)	bonsiep1	htm
WebSite Methodik - Vorlagen und Richtlinien für den Autor	autor2	htm

Abb. 2b: Bei nur 19 Dokumenten kann man sich noch vergleichsweise leicht einen inhaltlichen Überblick verschaffen, aber dennoch zeigt sich bereits der zusätzliche Orientierungseffekt durch die verdichtete deskriptorbasierte Darstellung in Abb. 2a.

Eine andere Interpretation faßt die generierten Deskriptorenlisten als Repräsentation einer dynamisch und nutzerangepaßte Hierarchie (Menüauswahl) auf. Jeder Deskriptor kann vom Nutzer an die Spitze der Hierarchie gewählt werden. Generiert wird die jeweils nächste Hierarchieebene, also die Menge der jeweils möglichen erfolgreichen Fortsetzungen einer einschränkenden Suche. Die Auswahlentscheidungen des Nutzers bestimmen also die spezifische, ihm angebotene Themenhierarchie vom Allgemeinen zum Speziellen.

Angestrebte Vorteile: Der wesentliche Vorteil der dargestellten Suchstrategie soll sein, daß eine hierarchische, zielgerichtete Suche unterstützt wird, ohne daß vordefinierte klassifikatorische Strukturen erforderlich sind, die leicht an der aktuellen Perspektive des Nutzers vorbeigehen können. Und außerdem soll der Nutzer, als Nebeneffekt der Suche, darüber informiert werden, was sonst er hätte alles finden können - und was nicht. Der Nutzer soll also eine Vorstellung davon aufbauen, inwieweit Themen (repräsentiert als Deskriptoren bzw. UND-

Verknüpfung von Deskriptoren) zentral oder peripher abgedeckt werden.

2.4 Gestaltungsentwurf

Das Thema "Retrieval"

Keywords: WebSite Methodik; Index; Suche; dynamisches Dokument

Das Thema "Retrieval" auf dem Server WebSite 'Methodik'



41 Dokumente zum Thema "Retrieval"

1. Beispiel (Klassifikationssystem): "Predicta Wörterbuch und Thesauri"
2. Beispiel (Klassifikationssystem): "RIS Classification" (D. Sengul, Okt. 1997)
3. Beispiel (Thesauri): "OED-12 Thesaurus (1993)"
4. Beispiel (Thesauri): "Alcohol and Other Drug (1992)"
5. Beispiel (Thesauri): "INPODATA: Thesaurus / Index (1994)"
6. Fachzeitschrift SIFT: Information Retrieval und Computerlinguistik
7. Internet-Quellen zum Thema "Information Retrieval Engines"
8. LDV Forum 97/1: "Inhaltsbasiertes Information Retrieval: Die Text Mining Technologie" (J. Gressel)
9. Leihkon-Sicherung: Information Retrieval Anwendungen (92)
10. Overhead-Folie: Das Rechner-Panorama
11. Overhead-Folie: Dokumente entstehen...
12. Overhead-Folie: Dokumentensystem wie Faktenretrieval
13. Overhead-Folie: Ein Dokument verändert sich...
14. Overhead-Folie: Entwicklung von Informationssystemen Overhead-Folie: Faktenretrieval "Holen statt suchen"
15. Overhead-Folie: Faktenretrieval: Archivische Fragestellung
16. Overhead-Folie: Offener Wörterbuch für Retrievalsysteme: TROG
17. Overhead-Folie: Qualitäts-Veränderung des Retrievalergebnisses...
18. Overhead-Folie: Relevanz-Feedback
19. Overhead-Folie: Retrieval von Bildern: Aufgabenklassifikation
20. Overhead-Folie: Retrieval von Bildern: Formähnlichkeiten
21. Overhead-Folie: Retrieval von Multimediale: Retrieval Wissen

Dokumente über "Retrieval"...

... thematisch unterteilt: ... nach Form unterteilt:

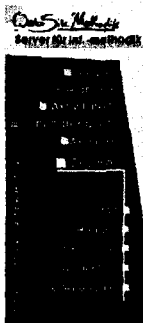
Anz. von Dokumenten über Retrieval UND ...	Anz. von Dokumenten über Retrieval als ...
11 Inhaltserschließung	21 Overhead-Folie
9 Automatische Indexierung	14 Beispiel
5 Evaluation	10 publiziertes Dokument (Dokumententyp)
Graphik	6 Skript
Diagramme	3 Übung
Informationslinguistik	2 Faksimile
Informationswissenschaft	Photo
Thesauri	
4 Dokumenten-Management	
3 Terminologie	
2 Externe Institution	
Klassifikationssystem	
Multidokumentation	
Ordnungssystem	
1 Datenbank	
Funktionsprototyp	
Klassifikation	
Internet	
Publikation	
Statistik	
Wiederholungsinformation	
Wörterbuchentwicklung	

Abb. 3: Exemplarische Darstellung eines Suchergebnisses am Beispiel „Retrieval“. Die s/w- Wiedergabe läßt die unterschiedliche Farbgestaltung von linker und rechter Bildschirmhälfte nur ansatzweise erkennen. Die Dokumentenliste zeigt, daß eine systematische Benennung der Dokumente den Überblick darüber unterstützt, welche Arten von Dokumente (Übung, Beispiel, Folie, Skript, etc.) vorliegen.

Es ist klar erkennbar, daß Retrieval auf diesem Server in erster Linie zusammen mit Inhaltserschließung/automatischer Indexierung behandelt wird.

Eine einfache Umsetzung des vorgestellten Konzeptes illustriert Abb. 3 an einem Beispiel. Ein in seiner Grundstruktur zweigeteilter Bildschirm bietet als Ergebnis der Auswahl des Deskriptors „Retrieval“ alle Dokumente zum Thema an, deren Durchsicht aber bereits längeres Scrollen erfordert. Parallel werden die zusammen mit „Retrieval“ indextierten Deskriptoren, nach inhaltlich und formal orientierten Deskriptoren unterschieden, aufgelistet. Die einzelnen Häufigkeitsränge werden durch Farbabstufung (in Gelb) so abgesetzt, daß jeder Rang (z.B. Rang 1 mit Inhaltserschließung, 11 mal gemeinsam mit Retrieval indexiert) gegenüber den nachfolgenden hervorgehoben erscheint.

Durch Auswahl eines der dargebotenen Deskriptoren (hier: Inhaltserschließung) wird die Suche fortgesetzt, was zu einem strukturell gleichartigen Ergebnis für das Thema „Retrieval und Inhaltserschließung“ führt (Abb. 4).



Das Thema "Retrieval" und "Inhaltserschließung" auf dem Server WebSite 'Methodik'



11 Dokumente zum Thema "Retrieval" und "Inhaltserschließung"

Dokumente über "Retrieval" und "Inhaltserschließung"...

1. Beispiel (Klassifikationssystem): "Prüfungsausschuss und Thesen"
2. Beispiel (Klassifikationssystem): "RIS Klassifikation" (D. Sauer, 1997)
3. Beispiel (Thesaurus): "QTD 12-Thesaurus (1985)"
4. Beispiel (Thesaurus): "Alcohol and Other Drug (1993)"
5. Beispiel (Thesaurus): "INFODATA-Thesaurus (1994)"
6. LIT-Forum 7/11: "Inhaltsverzeichnis Information Retrieval: Die Testungstechniken" (G. Gansel)
7. Ormsted-Forschung: Indizierung komplexer Dokumente
8. Stage: Indizes ... Kap. 2: Extraktionsmethode
9. Elvira: Indizes ... Kap. 2: Extraktionsmethode

... thematisch unterteilt:

... nach Form unterteilt:

Anz. von Dokumenten über "Retrieval" und "Inhaltserschließung" UND ...

Anz. von Dokumenten über "Retrieval" und "Inhaltserschließung" als ...

- 4 Thesaurus
- 3 Informationsquellen: Informationswissenschaft
- 2 Automatische Indizierung
- Erhebung
- Externe Indizes
- Klassifikationssystem
- 1 Publikation
- 1 Veranstaltung
- 1 Wechselseitige

- 7 Beispiel
- 4 publiziertes Dokument
- 3 Skript
- 2 Faksimile
- 1 Overhead-Folie

Abb. 4: Zur Fortsetzung der Suche (Abb. 3) wurde der Deskriptor Inhaltserschließung ausgewählt.

Ranking versus Ortskodierung: Der bisherige Vorschlag ist dadurch gekennzeichnet, daß er die häufig vorkommenden Themen an die ersten Positionen der Liste setzt (Ranking). Nur innerhalb eines Ranges kann etwa über eine Notation thematisch sortiert werden (was nur bei den kleinen Häufigkeiten einen Effekt hat, und auch dieser Effekt ist gering. Siehe dazu Abb. 2a).

Alternativ könnte auch eine Ortskodierung gewählt werden, bei der ein Deskriptor absolut oder relativ zu anderen immer an der gleichen Position einer Liste auftaucht.

Text versus Graphik: Alternativ zu der listenorientierten Oberfläche könnte (aufwendiger) auch eine graphische Darstellung etwa in Form eines Ausschnittes eines Begriffsnetzes gewählt werden oder als Ebene, aufgespannt vom Klassifikationsschema und der Kookkurrenzhäufigkeit, in der die einzelnen Deskriptoren eingetragen werden.

Einfachheit versus Funktionsvielfalt: Alternativ zur Beschränkung der Suchmethode auf eine einschränkende Suche (mittels UND-Operator) kann damit experimentiert werden, inwieweit auch Möglichkeiten anderer Suchstrategien (ODER-Operator, Einbeziehung begrifflicher Relationen) angeboten werden können, ohne das Interface und dessen Bedienung zu überfrachten.

3 Realisierung

3.1 Vokabular und Indexierung

Grundlage der Umsetzung des vorgestellten Konzeptes der thematischen Suche ist die Dokumentenbasis von WebSite 'Methodik, ein Vokabular von ca. 100 Deskriptoren sowie eine entsprechende Dokument-Indexierung.

Das Vokabular hat die Schwäche, daß es begleitend zum Aufbau des Informationsservers entstand. Die Entscheidung über das generische Niveau des Deskriptoren fiel auf der Basis der prognostizierten quantitativen Verhältnisse und unter dem Aspekt, Anpassungen bereits indexierter Dokumente zu vermeiden. In diesem Sinne sind einige (zentrale) Deskriptoren zu allgemein gewählt und die Unterbegriffe sind Nichtdeskriptoren. Es hat andererseits den Vorteil, daß die Deskriptoren sehr zielgenau der Logik des Faches und seiner Lehrveranstaltung entsprechen, so daß in diesem Zusammenhang sehr effektiv Dokumente aufgesucht werden können.

Jeder Deskriptor ist in ein Klassifikationssystem eingebunden, um fachsystematische Sortierungen von Deskriptormengen zu ermöglichen.

3.2 Datenbank

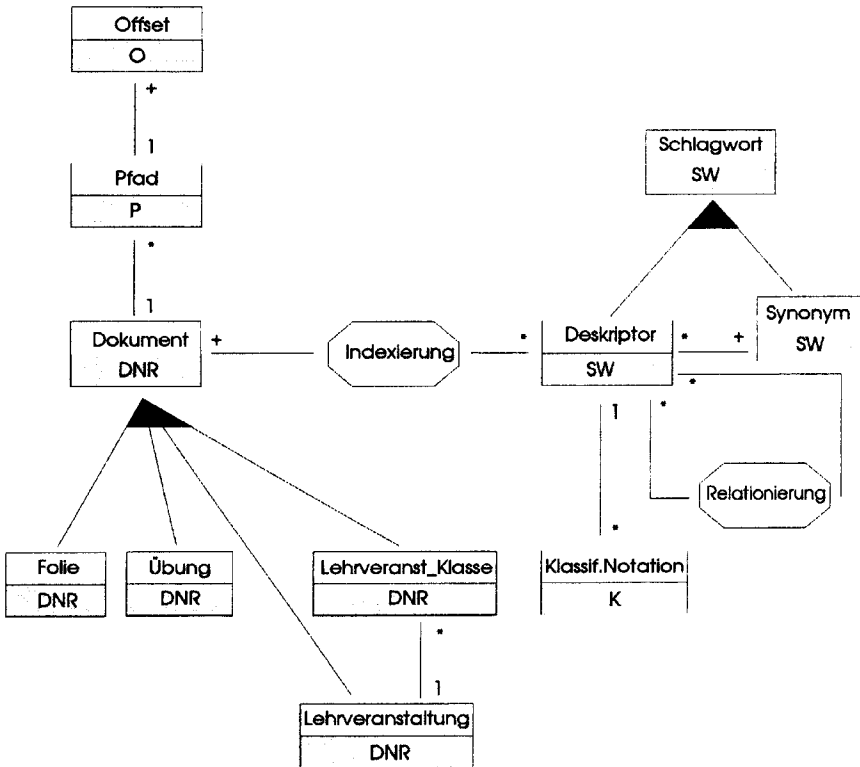


Abb. 5: Konzeptioneller Datenbankentwurf. Im Zentrum steht die Beziehung „Indexierung“ zwischen Dokumenten und Deskriptoren.

Die Datenbank stellt den Kern einer Datenhaltung dar, die in weiterer Zukunft immer mehr Aufgaben übernehmen soll. Ziel ist es, den Anteil von Informationen in statischen HTML-Dokumenten zurückzudrängen, um die entsprechen-

den Seiten dynamisch aus der Datenbank generieren zu können.

Den konzeptionellen Datenbankentwurf in einer Notation nach (Dennert 1991) zeigt Abb. 5. Die Umsetzung in Datenbankrelationen (unter Verwendung von ACCESS) erfolgt weitestgehend schematisch (Knorz 1996) und wird an dem SQL-Statement konkret, das die Tabelle in Abb. 2a erzeugt. Man findet in diesem Statement die Objektklassen des ER-Modells als Namen der Datenbankrelationen mit den entscheidenden Attributen wieder.

Obwohl der SELECT-Ausdruck 5 Relationen (die Relation Indexierung zweimal) in einem Join verbindet und eine Schachtelung enthält, ergeben sich keinerlei Effizienzprobleme.

```
SELECT Schlagwort.Bezeichnung,  
       count(Schlagwort.Bezeichnung) AS Anz, Klassifikation.Notation  
FROM Schlagwort, Indexierung AS I1, Indexierung AS I2, Deskriptor,  
       Klassifikation  
WHERE I1.SW = (SELECT Schlagwort.SW  
              FROM Schlagwort  
              WHERE Schlagwort.Bezeichnung="Gestaltung")  
AND I1.DNR=I2.DNR  
AND I2.SW=Schlagwort.SW  
AND I2.SW=Deskriptor.SW  
AND Deskriptor.K=Klassifikation.K  
AND Klassifikation.Notation like 'S*'  
  
GROUP BY Klassifikation.Notation, Schlagwort.Bezeichnung  
ORDER BY 2 DESC;
```

Abb. 6: SQL-Statement zur Erzeugung der Tabelle in Abb.2. Die letzte Selektion (like 'S') ist dafür zuständig, das Ergebnis anhand der Klassifikationsnotation auf inhaltlich orientierte Deskriptoren einzuschränken.*

3.3 Stand der Entwicklung

Im Rahmen einer Diplomarbeit am Fachbereich Information und Dokumentation ist das Konzept der thematischen Suche prototypisch realisiert worden (Windows NT Internet Server, Internet Database Connection und Access-Datenbank). Allerdings waren Datenbankstruktur und –zugriff für eine Verwendung im laufenden Betrieb nicht geeignet.

Die Abbildungen 3 und 4 sind nicht originale Bildschirmabzüge der vollzogenen Systemumstellung, sondern eine von mehreren Gestaltungsstudien, die mit realen Daten das Konzept illustrieren sollen und Vorgaben für die Implementierung machen.

Mittlerweile ist die Datenbank gemäß Abb. 5 vollständig aufgebaut. Alle Daten, soweit benötigt, stehen darin zur Verfügung, genauso wie die erforderlichen SQL-Abfragen. Es wird nun begonnen, das Konzept in seiner einfachsten Form (entsprechend Abb. 3 und 4) auf der Basis von ASP (Active Server Pages) für

Windows NT Internet Server umzusetzen und zur Verfügung zu stellen. Direkt anschließend ist eine Evaluierung geplant.

Literatur

[Bourdoncle 1997]

Bourdoncle, F.: *LiveTopics – recherche visuelle d'information sur l'internet*. In: Dossiers de l'Audiovisuel, La Documentation Française, (1997), 74, S. 36-38

[Beheim et al. 1998]

Beheim et al.: *Distance Learning for Students via Internet - a new Catchword for Studying?* Erscheint in: Shaping the Knowledge Society. Proceedings of BOBCATSSS 1998, Budapest, 1998.

Preprint: <http://www.iud.fh-darmstadt.de/iud/wwwmeth/paper/bobcat97/paper/vortrag/DL-Proceeding-HTML/DL-Proceeding-Final.html>

[Bürdek et al. 1998]

Bürdek, B.E.; M. Eibl; J. Krause: Visualisation in Document Retrieval: An example of the integration of software ergonomics and an aesthetic quality in design. <http://www.bonn.iz-soz.de/research/information/vri98/vri98eng.htm> (vorläufige Fassung; erscheint in den Proc. zu: Visual Representations and Interpretations 98, Liverpool, September 1998), 1998.

[Däßler/Otto 1996]

Däßler, R.; A. Otto: *Knowledge Browser - ein VRML-basiertes Navigations-tool für Information Retrieval Systeme im World Wide Web*. Internationales Symposium für Informationswissenschaft (Proc. ISI'96, Berlin), Konstanz: Universitätsverlag Konstanz, 1996, S. ????

[Däßler/Palm 1998]

Däßler, R.; H. Palm: *Virtuelle Informationsräume mit VRML . Information recherchieren und präsentieren in 3D*. Heidelberg: dpunkt, 1998

[Dennert 1991]

Denert, E.: *Software Engineering*. Berlin, Heidelberg et al.: Springer, 1991.

[Frisch/Kluck 1997]

Frisch, E.; M. Kluck: *Pretest zum Projekt German Indexing and Retrieval Testdatabase (GIRT) unter Anwendung der Retrievalsysteme Messenger und freeWAISsf*. Informationszentrum Sozialwissenschaften Bonn, IZ-Arbeitsbericht Nr. 10, 1997

[Hemmje 1993]

Hemmje, M.: *Eine inhaltsorientierte, intuitive 3D-Benutzerschnittstelle für Information-Retrieval-Systeme*. In: Knorz, G., Kraus, J. & Womser-Hacker, C. (Hr.): *Information Retrieval '93. Von der Modellierung zur Anwendung* (Proc. IR'93, Regensburg), Konstanz: Universitätsverlag Konstanz, 1993, S. 82-99

[HIS 1998]

HIS: *Dokumentation medienunterstützten Lehrens und Lernens*.

HIS: Hochschul-Informations-System-GmbH. (Recherchierbare WWW-Adressen-Sammlung zum Einsatz elektronischer Medien an Hochschulen).

<http://www.his.de/abt3/proj/676/index.html> (20.8.1998).

[Kleinberg 1998]

Kleinberg, J.: *Authoritative Sources in a Hyperlinked Environment*.
Proceedings of the 9th ACM-SIAM Symposium on Discrete Algorithms, 1998

[Knorz 1996]

Knorz, G.: *Datenbank-Entwurfsmethoden*. In: Buder, M; Rehfeld, W.; Seeger, Th. (Hr.): *Grundlagen der praktischen Information und Dokumentation*. Bd. 2., 4. Ausgabe. München, London, New York, Paris: K.G. Saur, 1996, S. 664-687

[Knorz 1998]

Knorz, G.: *WebSite ,Methodik als zentrale Informationsressource für Lehren und Lernen*. LDV-Forum 15(1998), 1, S. 45-58

Preprint: <http://www.iud.fh-darmstadt.de/iud/wwwmeth/paper/ldvf97b/paper1.htm>

[Kollewe et al. 1994]

Kollewe, W.; M. Skorsky; F. Vogt und R. Wille: *TOSCANA - ein Werkzeug zur begrifflichen Analyse und Erkundung von Daten*. In: R. Wille; M. Zickwolff (Hr.), *Begriffliche Wissensverarbeitung - Grundfragen und Aufgaben*, B.-I. Wissenschaftsverlag, Mannheim-Leipzig-Wien-Zürich, 1994, S. 267-288.

[Paris/Tibbo 1998]

Paris, L.A.H.; H.R. Tibbo: *Freestyle vs. Boolean: A Comparison of Partial and Exact Match Retrieval Systems*. *Information Processing and Management*, 34(1998),2,3, S. 175-190

[Rorvig 1998]

Rorvig, M.: *Scaled Structure in Visualized TREC Data and Query Feedback*. *Information Processing and Management*, 34(1998),2,3: S. 151-160

[Spoerri 1994]

Spoerri, A.: *InfoCrystal: A Visual Tool for Information Retrieval & Management*. In: M. Galer; S. Harker; J. Ziegler (Hr.): *Conference on Human Factors in Computing Systems (Proc. ACM CHI'94)*, 1994, S. 11-12

[Stein 1997]

Stein, P.: *Verbesserung von Suchindizes im World Wide Web*. Ludwig-Maximilians-Universität München, CIS-Bericht, 1997, S. 97 - 105

Online-Marketing von Rundfunkmedien – Ergebnisse einer empirischen Studie

Claus Rautenstrauch, Jochen A. Siegle

Otto-von-Guericke-Universität Magdeburg
Institut für Technische und Betriebliche Informationssysteme (FIN/ITI)
Postfach 4120
D-39016 Magdeburg
Email: rauten@iti.cs.uni-magdeburg.de

Inhalt

- 1 Ausgangssituation
- 2 Grundlage der Untersuchung
- 3 Theoretische Vorbemerkungen
- 4 Institutionelle Klassifizierung
- 5 Bestückungskomponenten von Information-Sites
 - 5.1 Online-Verkauf
 - 5.2 Online-Direct-Marketing/Online-Direct-Mail
 - 5.3 Online-Marketingforschung
 - 5.3.1 Online-Marketingforschung von Rundfunkanstalten
 - 5.3.2 Ergebnisse
- 6 Evaluierung der Online-Angebote von Rundfunkanstalten
 - 6.1 Erfolgsfaktor Interaktivität
 - 6.2 Erfolgsfaktor Mehrwerte
 - 6.3 Sonstige Erfolgsfaktoren

Zusammenfassung

In diesem Beitrag werden Ergebnisse einer empirischen Untersuchung zur Präsenz von Rundfunkanstalten in World Wide Web (WWW) vorgestellt. Es wird gezeigt, in welcher Weise Rundfunkanstalten das WWW nutzen. Nach einer kurzen Einführung in die Thematik wird gezeigt, welche Basisformen des Online-Marketing in welcher Weise unterstützt und welche relevanten Erfolgsfaktoren in welcher Weise durch den WWW-Auftritt erreicht werden. Den Abschluß bildet eine Bewertung der präsentierten Ergebnisse.

1 Ausgangssituation

Nach Radio und Fernsehen ist das World Wide Web zum weiteren Betätigungsfeld von Rundfunkanstalten geworden, auf dem Informationen und Unterhaltung verbreitet werden. Die Web-Auftritte der Rundfunkanstalten sind durchweg aufwendig und professionell. Insbesondere bei der Nutzung multimedialer Elemente auf WWW-Seiten sind die Rundfunkanstalten im Vergleich zu anderen Anbietern weit vorne dabei. So wurde auch ernsthaft seitens der öffentlich-rechtlichen Rundfunkvertreter vorgeschlagen, die Rundfunkgebühren auch auf die Nutzung des World Wide Web auszudehnen, was – glücklicherweise – von der Politik jedoch bislang abgelehnt wird. Die Hintertür steht jedoch offen, da insbesondere für private Haushalte sogenannte Web-TV-Geräte („Set-Top-Boxen“), die eine Kombination aus WWW-Technologie und Fernsehen darstellen, vor der Markteinführung stehen. Im Internet-Mutterland USA wurde bereits fast eine halbe Million derartiger Empfangsgeräte verkauft, in Deutschland soll den Geräten mit dem Weihnachtsgeschäft 1999 der breite Durchbruch gelingen. Über die Web-TV-Technologie steht das World Wide Web auf der Schwelle zum Massenmarkt der privaten Haushalte, und die Rundfunkanstalten haben beachtlichen Anteil daran. Besonders verdeutlichen dies die monatlich ausgewiesenen Reichweitenzahlen der Online-Medien: Online-Angebote von Rundfunkanstalten verzeichnen neben den Angeboten von Printmedien die meisten monatlichen Zugriffe¹.

Diese Entwicklung fand jedoch insgesamt unter weitgehendem Ausschluß der Wissenschaft statt, von daher wird in diesem Beitrag eine systematische Analyse der Web-Auftritte öffentlich-rechtlicher und privater Rundfunkanstalten vorgestellt und werden Entwicklungspotentiale aufgezeigt.

Ziel der Untersuchung ist, ausgehend von einer präzisen Erarbeitung eines Online-Marketing-Konzeptes, zu untersuchen, inwieweit und in welchen Ausprägungen WWW-Auftritte als Form der kommerziellen Kommunikation für Rundfunkanstalten Relevanz hat. Dabei ergibt sich aus der Untersuchung ein umfassender Überblick über die Online-Marketing-Praxis bzw. -Aktivitäten von deutschen Rundfunkanstalten.

2 Grundlage der Untersuchung

Die Online-Angebote wurden im Hinblick auf die relevanten Online-Marketing-Erfolgsfaktoren (z.B. Interaktivität, Non-Linearität, Multimedialität, Serviceorientierung etc.) anhand einer im Zeitraum von Mai–Juni 1997 durchgeführten quantitativen Untersuchung der Web-Angebote deutscher TV- und Radio-Stationen evaluiert. Als Panel für diese Erhebung wurden nach folgenden Kriterien ausgewählt:

¹ ausgewiesen durch die IVW (Informationsgesellschaft zur Feststellung der Verbreitung von Werbeträgern) in „PageImpressions“ oder „Visits“ (vgl. <http://www.ivw.de>)

- alle öffentlich-rechtlichen Rundfunkanstalten mit Online-Angebot im WWW
- alle national sendenden privaten Fernsehanstalten mit Online-Angebot im WWW
- alle privaten Hörfunkanstalten mit einer Gesamtreichweite von über 100.000 Hörern² und Online-Angebot im World Wide Web

Gemäß diesen Auswahlkriterien umfaßt das Panel – kurz „RFA-Panel“ – 41 Online-Angebote von Rundfunkveranstaltern. Diese lassen sich wiederum in 23 öffentlich-rechtliche und 18 private Online-Angebote von Rundfunkveranstaltern klassifizieren. Die Urliste zur Erhebung ist im World Wide Web unter <http://www.wildersueden.de/jasiegle.htm> abgelegt. Eine ausführliche Darstellung der Ergebnisse in (Siegle 1997) dokumentiert.

3 Theoretische Vorbemerkungen

Gemäß (Fantapié Altobelli/Hoffmann 1996) existieren in Online-Diensten als Basis-Online-Marketingformen die Information-Site, Online-Sponsoring und Online-Shopping. Um sich der Frage zu nähern, inwieweit Rundfunkanstalten generell Online-Marketing betreiben, ist zunächst darauf hinzuweisen, daß diese in keiner erwähnenswerten Form klassische Online-Werbung in Form von Online-Sponsoring betreiben. Zumindest derzeit plazierte keine deutsche Rundfunkanstalt Buttons oder Banner mit oder ohne Hyperlink zur eigenen Homepage auf fremden Internet-Seiten. Eine Ausnahme bildet hierbei der digitale Kirch-Kanal *DF1*, welcher innerhalb des Online-Angebotes „VH1derland“ vom Musiksender *VH1* einen aktiven Banner eingeschaltet hat. Da keine anderen Rundfunkanstalten Online-Sponsoring betreiben, ist davon auszugehen, daß diese Basis-Online-Marketingform nicht Teil des Marketing-Mixes der Sender ist. Online-Sponsoring spielt allerdings für Rundfunkanstalten in der Form eine Rolle, daß sie zur (Re-)Finanzierung von Online-Angeboten der Rundfunkveranstalter beiträgt.

Die zweite Online-Basismarketingform Online-Shopping bezieht sich nach *Fantapié Altobelli* auf den „Verkauf von Produkten über einen virtuellen Marktplatz (Shopping Mall) im Internet. Dargestellte Produkte können einen Hyperlink zur eigenen Information-Site beinhalten.“ (Fantapié Altobelli/Hoffmann 1996, S.26)

Definitionsgemäß trennt *Fantapié Altobelli* das Online-Shopping, d.h. den POS (Point of Sale), geographisch von den Information-Sites. Da Rundfunkanstalten – sofern überhaupt – Online-Shopping nur innerhalb ihrer eigenen Online-

² Reichweiten-Angaben ausgewiesen nach dem „Funk-Planungsservice 1997“ der RMS – Radio Marketing Service GmbH, Hamburg 1997

Angebote betreiben, schließt die Begriffsdefinition die Online-Basismarketingform Online-Shopping für Rundfunkanstalten aus. Da verschiedene Rundfunkanstalten jedoch rudimentäre Formen von Online-Shopping – meist in Form von Merchandising-Angeboten – innerhalb ihrer Information-Sites betreiben, wird dieser Bereich nicht gänzlich ausgeklammert, sondern unter den Bestückungskomponenten der Information-Sites von Rundfunkanstalten als „Online-Verkauf“ i.S.v. *Fantapié Altobelli* aufgegriffen.

Unter der dritten Online-Basismarketingform, der Information-Site, ist „der Auftritt eines Unternehmens im WWW mit eigener Internetadresse bzw. -Site zu verstehen“ (*Fantapié Altobelli/Hoffmann 1996, S.23*), was bereits durch die Präsenz mit einem Online-Angebot einer Rundfunkanstalt gegeben ist. Demgemäß ist die einzige Basisform, mit der Rundfunkanstalten Online-Marketing betreiben, das Online-Marketing über eine Information-Site.

Daher wird ausführlich auf die einzelnen Information-Sites bzw. Online-Auftritte von – exemplarisch ausgewählten – Rundfunkanstalten eingegangen. Hierbei scheint eine Kategorisierung von privaten und öffentlich-rechtlichen Rundfunkveranstaltern sinnvoll, da sich anhand dieser strukturellen Unterscheidung verschiedene Bedingungen für die jeweiligen Online-Angebote ergeben. Eine Unterscheidung von Hör- und Fernsehfunk erscheint hier nicht notwendig, da es, wie gezeigt wird, zunächst für ein WWW-Angebot keinen Unterschied macht, ob die Inhalte von einer Radio- oder TV-Station oder einer Kombination beider erbracht werden. Der abstrakte Begriff der Information-Site wird im folgenden synonym mit den Begriffen Online-Angebot bzw. Online-Auftritt verwendet.

4 Institutionelle Klassifizierung

Hinsichtlich des institutionellen Bezugs der Online-Angebote von Rundfunkanstalten klassifizieren *Schröter* und *Ewald* drei Typen (*Schröter/Ewald 1996, S. 478*): gesamthausbezogene, programmbezogene und sendungsbezogene Web-Angebote. Im folgenden werden die Online-Angebote der Rundfunkanstalten des RFA-Panels dieser Kategorisierung folgend klassifiziert.

Etwaige Zuordnungen zu den Typen gesamthaus- bzw. programmbezogenes Angebot verschwimmen bei verschiedenen Sendern – vor allem bei denjenigen mit genau einem Programm. Sendungsbezogene Online-Angebote beschäftigen sich inhaltlich mit nur einer Sendung.

Die institutionelle Kategorisierung macht verschiedene Mehrfachnennungen notwendig, so z.B. bei SAT.1. Der Mainzer Privatsender erweiterte sein bis Sommer 1996 nur auf die „Harald-Schmidt-Show“ bezogenes – daher rein sendungsbezogenes – Online-Engagement um ein gesamthausbezogenes. Das Online-Angebot zur Late-Night-Sendung wird allerdings weiterhin unter der Show-eigenen Domain präsentiert und vermarktet.

**GESAMTHAUS-
BEZOGEN****ÖFFENTLICH-RECHTLICHE:**

ARD, Westdeutscher Rundfunk (WDR), Mitteldeutscher Rundfunk (MDR), Bayerischer Rundfunk (BR), Norddeutscher Rundfunk (NDR), Süddeutscher Rundfunk (SDR), Hessischer Rundfunk (HR), Radio Bremen (RB), Saarländischer Rundfunk (SR), Deutsche Welle (DW), Deutschland-Radio/Deutschland-Funk (DLF/DLR), Sender Freies Berlin (SFB), Zweites Deutsches Fernsehen (ZDF), Arte, 3Sat

PRIVATE:

Pro Sieben, Kabel 1, RTL, RTL 2, SAT 1, VH1

**PROGRAMMBE-
ZOGEN****ÖFFENTLICH-RECHTLICHE:**

SWF3, ORF (Online-Angebot nur zu Radio Fritz), BR5 Aktuell

PRIVATE:

Hit-Radio FFH, Radio FFN, 104.6 RTL, 94.3 rs.2, Antenne Mecklenburg-Vorpommern, Antenne Das Radio, Hundert.6, Radio PSR, Radio Regenbogen, Radio RPR, Radio SAW, RSH, RTL-Radio, VOX, DF1

**SENDUNGSBE-
ZOGEN****PRIVATE:**

N-TV (Mouse & Co.), SAT 1 (Harald Schmidt Show, Ran)

5 Bestückungskomponenten von Information-Sites

Die Basis-Marketingformen können mit weiteren Marketingkomponenten bestückt werden. Von den mit *Fantapié Altobelli* beschriebenen Bestückungskomponenten sind als für Rundfunkanstalten im allgemeinen – vor allem aber auch für die dem RFA-Panel zugehörigen – nicht relevante Komponenten Online-Services, Database-Marketing, Online-Distribution, virtuelle Messen und Data-Warehouse zu kennzeichnen. Als Marketingkomponenten der Information-Sites von Rundfunkanstalten kommen dagegen der Online-Verkauf, die Online-Marketingforschung und Online-Direct-Marketing/Online-Direct-Mail in Be-

tracht. Die Bestückungskomponente „Electronic Publishing“ ist nur peripher relevant und wird hier daher nicht weiter behandelt.

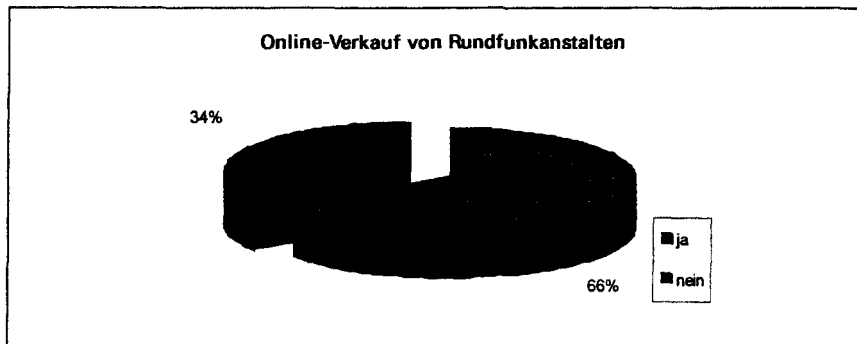
5.1 Online-Verkauf

Wie bereits in Kapitel 4 erwähnt, betreiben Rundfunkanstalten kein Online-Shopping im eigentlichen, beschriebenen Sinn. Die Sender beschränken ihre Online-Distribution i.d.R. fast ausschließlich auf den – nicht unlukrativen – Verkauf von Merchandisingartikeln. Im Jahr 1996 sollen alleine mit dem TV-Merchandising durch den Verkauf von Fan-Shirts, Baseballmützen, Büchern zum Film, CDs zu Serien etc. Handelsumsätze in Höhe von sechs Milliarden DM erwirtschaftet worden sein (Südkurier vom 3.4.1996, S.16).

Ebenso heterogen wie die „Online-Warenkörbe“ der Sender sind auch die Kaufabwicklungen bzw. Zahlungsbedingungen. So läßt bspw. *Radio Schleswig-Holstein* im *R.SH-Shop* direkte Online-Bestellungen von seinem sehr breit angelegten Merchandising-Programm zu. Und auch im *RTL 2-Internet-Shop* lassen sich – neben Büchern, CD-Eigenproduktionen und Fan-Shirts – in Kooperation mit Hardware-Häusern sogar ganze PC-Komplettsysteme direkt online bestellen.

Bei *KABEL 1* ist man wesentlich vorsichtiger: Artikel im Online 1-Store sind nur per Fax oder telefonisch bestellbar. Auch in der Online-Boutique beim grenzübergreifenden Kulturkanal *Arte* muß für eine Bestellung, z.B. von *Arte-Feuerzeugen* oder *-Seesäcken*, ein Formular ausgedruckt und per Fax oder „Snail-Mail“ versandt werden. Schon fast zu restriktiv sind die Konditionen für einen Online-Einkauf beim *SFB*, denn aus der Hauptstadt wird nichts ausgeliefert, bevor nicht ein Verrechnungs-Scheck an die *SFB-Werbung* geschickt oder der fällige Betrag auf ein *SFB-Konto* überwiesen wurde. Im sehr aufwendig gestalteten *PRO SIEBEN-Shop* kann dagegen sowieso nur einkaufen, wer *PRO SIEBEN-Club-Mitglied* ist.

ABB. 1: ONLINE-VERKAUF BEI ONLINE-ANGEBOTEN VON RUNDFUNKANSTALTEN



Wie hoch die Anteile der Umsätze aus Online-Shopping an den Gesamt-Merchandising-Umsätzen der Sender sind, kann nur geschätzt werden. Konkrete Daten hierzu liegen nur aus dem Ausland vor, und zwar vom *Schweizer Fernsehen SF DRS* (www.sfdrs.ch). Gemäß Joachim Frey, Verantwortlicher für das Online-Angebot beim *Schweizer Fernsehen*, zählte man von Anfang Dezember 1996 bis Mai 1997 22.000 Zugriffe auf die WWW-Seiten des „TV-Shops“; Online-Bestellungen gingen in diesem Zeitraum etwa 350 ein.

Verstärkte Bestrebungen, den Online-Shopping-Bereich weiter zu professionalisieren, zeigt der *SFB*. Auch wenn derzeit keine Informationen über die Pläne für kommenden Herbst nach außen dringen, hat man immerhin schon eine eigene Domain für den *SFB-Shop* www.sfb-shop.de reservieren lassen.

Einen Vorstoß in bezug auf Online-Shopping möchte auch Nachrichtensender *n-tv* starten und präsentiert deshalb auf der *Internet World Spring '97* im Juni 1997 in München den ersten Online-Shop mit TV-Anbindung. Auch wie dieser sich präsentieren wird, ist derzeit noch Spekulation.

Zusammen mit dem *Nomos-Verlag*, *IBM* und *ARS* startete der *Südwestfunk* im Mai 1997 das Einkaufszentrum „TV-Markt.de“ (www.tv-markt.de). Ergänzt durch einen Computershop von *IBM* und *ARS*, gibt es dort Bücher von *Nomos* und *SWF-Merchandising*-Artikel zu kaufen. Als Standardsoftware wurde die *IBM*-eigene „net.commerce“ verwendet. Das Projekt der *SN Südwest-Net* – ein Gemeinschaftsunternehmen der *Südwestfunk Media* und dem *Nomos-Verlag*, will mittelständischen Unternehmen eine einfache Möglichkeit für die Ausdehnung von Verkaufsaktivitäten auf das Internet bieten.

5.2 Online-Direct-Marketing/Online-Direct-Mail

Derzeit ist in Deutschland kein Sender für die Praktik bekannt, unaufgefordert individualisierte Massen-Mails zu versenden. Dagegen machen sich verschiedene Anstalten das Online-Marketinginstrument „E-Mail-Response-Services“ zunutze, um persönlichen Kontakt zu ihren Zuhörern bzw. -schauern zu halten. So setzen bspw. *SAT.1* und der Radiosender *Lippewelle Hamm* ganz gezielt diese sehr kostengünstige Online-Direktmarketingform zur aktiven Kundenbindung ein und versenden regelmäßig Newsletters an einmal registrierte User – die Kosten für Mailinglisten werden inklusive der Ausgaben für das Handling auf etwa 10 Cent pro Adresse beziffert (*ZD Internet Professionell 4/97*, S. 50). Durch aktuelle Hinweise geben die weitestgehend automatisiert versandten Newsletters immer wieder den Anstoß, den Server aufzusuchen. Dabei ist vor allem zu beachten, daß der E-Mail-Response-Service in keiner Weise die „Netiquette“ verletzt, da die E-Mail-Informationen bestellt sind und nicht unaufgefordert ins Haus kommen (Hünerberg/Heise/Mann 1996, S. 187).

Neben der aktiven Kundenbindung dient den Rundfunkveranstaltern die Interaktion mit Hörern bzw. Zuschauern via E-Mail als Feedbackkanal und -kontrolle. Dieser Erkenntnis tragen alle Rundfunkanstalten aus dem RFA-Panel Rech-

nung und stellen eine E-Mail-Möglichkeit für User innerhalb ihrer Online-Angebote bereit.

Wie hoch die Akzeptanz der elektronischen Post als Feedbackkanal für Zuschauer/Hörer ist, demonstriert z.B. die Sendung „ZAK“. Der intellektuelle Wochendurchblick des WDR erhielt pro Sendung rund 300 Mails – ein Vielfaches dessen, was die Redakteure an Feedback gewohnt waren, bevor die E-Mail-Adresse in der Sendung eingeblendet wurde (PI@net 7/96, S. 31). Um den Ansturm zu bewältigen, wurden drei Mitarbeiter mit der Pflege der Seiten und der Mails betraut.

Damit trug der WDR der Forderung Rechnung, daß Online-Kommunikation beschleunigte Reaktionszeiten erfordert – nach Angaben des Fachmagazins „Horizont“ erwarten die Absender einer E-Mail innerhalb von 24 Stunden eine Antwort (Horizont 4/97, S. 48). Dieser essentielle Bestandteil eines Online-Auftritts wird „Feedback-Management“ genannt. Horizont.net – der Online-Dienst von „Horizont“ – testete im Januar 1997 den E-Mail-Service von 45 Unternehmen, Agenturen und Medienvertretern. Im Rahmen dieser Untersuchung antworteten von den insgesamt acht Medienvertretern SAT.1, TV Today Online und Die Zeit Online binnen 24 Stunden, Spiegel-Online nach zwei Tagen, PRO SIEBEN nach fünf Tagen und RTL, FAZ Online und Focus Online nach sieben Tagen. Obwohl zwar alle Medienvertreter antworteten, war dieses Ergebnis für die „Horizont“-Redaktion doch Anlaß genug, mit „der Service läßt zu wünschen übrig“ zu bewerten.

Die im Rahmen dieser Arbeit durchgeführten eigenen Recherchen können dies zum Teil bestätigen: E-Mails an den MDR, den SR sowie den Webmeister des ZDF blieben gänzlich unbeantwortet, und auch die Radio-Stationen Hundert,6 und RTL 104,6 zeigten keinerlei Reaktion. Bei ARTE folgte immerhin auf den dritten – etwas unfreundlicheren – Mail-Versuch ein Anruf. Am schnellsten antworteten die ARD und RTL 2. Beide Anstalten senden direkt nach dem Eintreffen einer Message auf ihrem Server eine automatisierte – wenn auch nicht personalisierte – Rückantwort an den Adressaten der Mitteilung, bedanken sich für die Mail und versichern, die Anfrage zügig zu bearbeiten. Auch SAT.1 antwortet prompt auf jeden neu registrierten SAT.1-Online-User per E-Mail, zum einen, um das vereinbarte Paßwort zu bestätigen, zum anderen, um nochmals auf die Extra-Leistungen (z.B. das „Play & Win“-Spiel) hinzuweisen, welche registrierte Anwender in Anspruch nehmen können. In der „Horizont“-Untersuchung demonstrierte die Mainzer TV-Anstalt allerdings, daß trotz rascher Antwort Sympathie-Verluste bei Nutzern erfolgen können: Der Sender antwortete zwar innerhalb eines Tages, vertröstete aber per Standard-E-Mail den Empfänger darauf, die gewünschten Informationen in den nächsten Tagen per „Snail-Mail“ zu schicken – die erhoffte Antwort kam in der „Horizont“-Redaktion allerdings nie an.

5.3 Online-Marketingforschung

5.3.1 Online-Marketingforschung von Rundfunkanstalten

Online-Marketingforschung betreiben generell der *SWF*, der *WDR*, der *MDR*, *RTL*, *RTL 2* sowie *KABEL 1* und *PRO SIEBEN* über ihren Vermarkter, die *MediaGruppe München*. In welcher Form – aktiv/passiv und quantitativ/qualitativ – diese Anstalten Online-Forschung allerdings ausüben, ist nicht bekannt. Es kann jedoch davon ausgegangen werden, daß die Aussage „Online-Forschung zu betreiben“ hinreichend für die Annahme ist, daß die passive Form durch Auswertung von Log-Files angewandt wird.

An qualitativen Untersuchungen von Rundfunkanstalten ist eine Befragung von 896 Nutzern bekannt, die von *PRO SIEBEN-Online* im Zeitraum von August bis September 1996 durchgeführt wurde (MediaGruppe München (MGM) unter www.mgmuc.de/pro7-onl/einstieg.htm). Auch *RTL* gibt an, qualitative Methoden anzuwenden, um das Online-Angebot besser steuern und kontrollieren zu können. Das *ARD-Online-Angebot* zu „Das Erste“ beteiligt sich an der Hamburger „W3B-Studie“. Bei unseren österreichischen Nachbarn vom *ORF* betreibt die hauseigene Medienforschung „sehr intensive Internetforschung“. Seit einem halben Jahr werden PC-Besitz, Internetanschluß sowie Verweildauer vor dem Computer abgefragt; außerdem entwickeln die Wiener derzeit gemeinsam mit Medienforschern aus den USA ein objektives Meßsystem für das Internet.

Derzeit in Planung ist der Einsatz von Online-Forschungsinstrumenten bei *Radio RS.2*, *Viva* und dem *SFB*. Letzterer plant, im Rahmen seiner neuen Projekte, qualitative und quantitative Instrumente zu installieren. Definitiv keine Online-Marketingforschung betreiben momentan der *BR*, *RB*, *RSH* und der *ORB* mit *FRITZ*. Keine Angaben zu Online-Forschung machten dagegen der *NDR* und der *HR*.

Das aktive Online-Marketingforschungsinstrument der Anwenderregistrierung, bei welcher sich User vor jedem Abruf mit Namen und Paßwort identifizieren und meist einen Fragebogen zur Person ausfüllen müssen (Hünerberg/Heise/Mann 1996, S. 207), wird derzeit bei fünf Rundfunkanstalten aus dem RFA-Panel eingesetzt: bei *SAT.1*, *Hit Radio FFH*, *RADIO FFN*, dem *NDR* für die Teilnahme am „N-Joy-Chat“ und bei *DF1* für eine Selektionsanfrage an den „Electronic Programm Guide“. Alle anderen Rundfunkanstalten aus dem RFA-PANEL kommen ohne Anwenderregistrierung aus. Momentan generiert auch kein Sender aus dem Untersuchungspanel User-Daten über eine andere Online-Befragung.

SAT.1 generierte über die Anwenderregistrierung bereits die persönlichen Daten von 10.000 eingetragenen Mitgliedern. Für Inanspruchnahme von Leistungen aus dem Online-Angebot von *Hit Radio FFH* ist eine Registrierung beim – kostenlosen – *FFH-Provider Germany.net* (www.germany.net) notwendig. *Radio SAW* registriert zumindest den Namen der Surfer zwecks persönlicherer Ansprache beim virtuellen Rundgang durchs Funkhaus.

Als weiteres Instrument zur Adreßdatengenerierung dienen interaktive Online-Gewinnspiele, die derzeit von 20 Sendern aus dem RFA-Panel einsetzen.

Inwieweit Rundfunkanstalten qualitative Online-Forschung in Form von Auswertungen der extrem feinfühligten Feedbackmechanismen E-Mail und Gästebuch anwenden, ist nicht bekannt. Gästebücher offerieren allerdings 23 der 41 Untersuchungseinheiten des RFA-Panels; über den Feedbackkanal E-Mail verfügen alle Anstalten des Panels.

ABB. 2: ANWENDERREGISTRIERUNG FÜR INANSPRUCHNAHME BESTIMMTER LEISTUNGEN INNERHALB DES ONLINE-ANGEBOTES

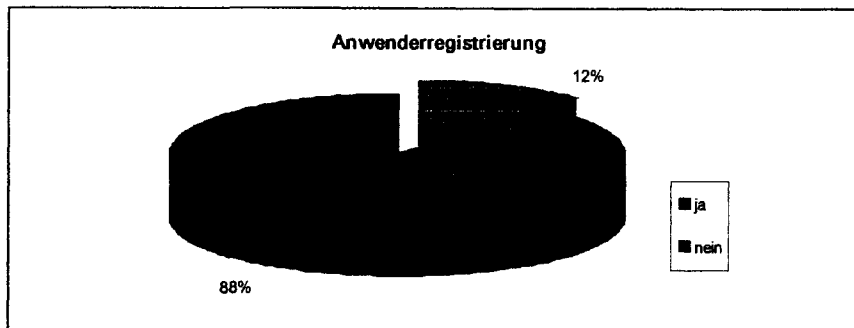
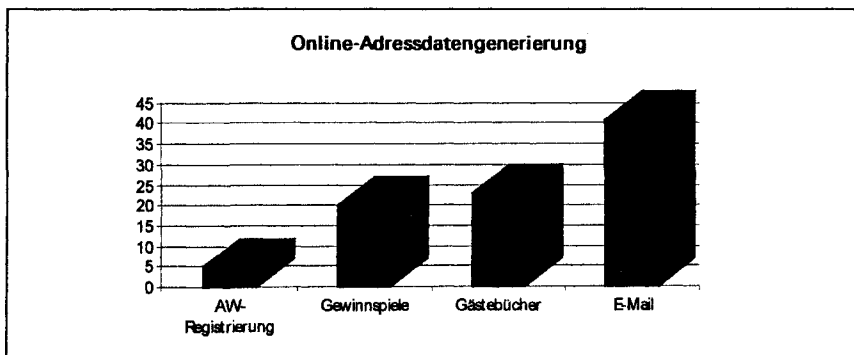


ABB. 3: ONLINE-ADRESSDATENGENERIERUNG



5.3.2 Ergebnisse

Im folgenden werden auszugsweise und exemplarisch verschiedene elementare Forschungsergebnisse zum Online-Angebot ausgewählter Rundfunkanstalten – zum einen bezüglich Nutzerdaten, zum anderen bezüglich Reichweiten der Online-Angebote – dargestellt.

Gruppe München (MGM) unter www.mgmuc.de/pro7-onl/einstieg.htm). Nachfolgend werden die wichtigsten Ergebnisse und Userprofile dieser Studie dargestellt werden. Des weiteren werden Auswertungen der ersten MDR-Nutzeranalyse hinsichtlich der Frage, über welche (Zugangs-)Server sich die User zu *MDR-Online* einloggen, angeführt.

Die erste Nutzeranalyse des *Mitteldeutschen Rundfunks* zeigt, woher die User des *MDR-Online*-Angebotes kommen: 39 Prozent über Server der Universitäten, sieben Prozent über Hoch- und Fachschulen, 16 Prozent nutzen *T-Online*, zwei Prozent *America Online* ein Prozent *CompuServe* und 36 Prozent diverse andere Server (Medien Bulletin 10/96, S. 26).

Zur Nationalität der User macht die *MDR*-Nutzeranalyse keine Angaben. Der *Bayrische Rundfunk* beschäftigt sich aber umso mehr mit der Frage, wie hoch der Anteil ausländischer Zugriffe auf das „Personalradio“-Angebot ist: 39 Prozent der *BR-Online*-User kommen aus dem Ausland, davon wiederum 31 Prozent aus den USA (Pressemitteilung WDR-Online 18.4.1997).

Abb. 4: INTERNET-ZUGANGS-PLATTFORMEN DER MDR-ONLINE-USER

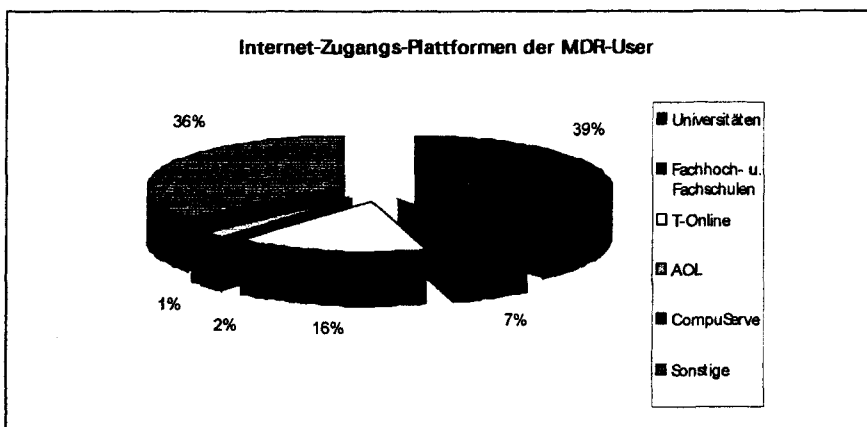


ABB. 5: EINKOMMENSSTRUKTUR DER PRO SIEBEN-ONLINE-USER

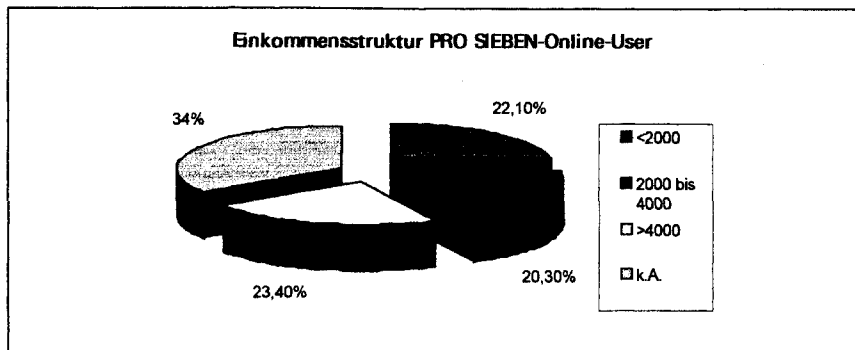
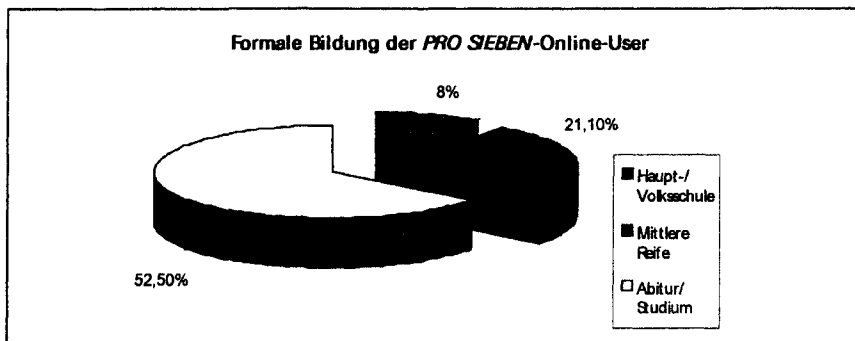


ABB. 6: FORMALE BILDUNG DER PRO SIEBEN-ONLINE-USER



TAB. 1: BEISPIELE FÜR REICHWEITEN DER ONLINE-ANGEBOTE VON RUNDFUNKANSTALTEN

SENDER	PAGE VIEWS 1997	PAGE VIEWS 1996
KABEL 1	600.000 (Jan.)	—
RSH	15.000 (Mai)	—
RS2	30.000 (Mai)	—
SAT 1	600.000 – 900.000 (Mai)	—
SAT 1 (H.-SCHMIDT-SHOW)	3.000.000 (Mai)	—
RTL 2*	87.000 (Jan.)	75.000 (Dez.)

Quelle: eigene Befragung 3–6/97

* IPA-Plus-Mappe zu RTL bzw. RTL 2. Die ausgewiesenen Zahlen berücksichtigen nicht die Zugriffe über Proxy-Server, d.h. die realen Zugriffszahlen liegen weitaus höher.

** Daten aus PRO SIEBEN-Logfile-Auswertung. Basis: Vollständige HTML-Seiten (ohne Chat, eigene Zugriffe, Frames, Client Push). Vgl. MediaGruppe München (MGM) unter www.mgmuc.de/pro7-on/einstieg.htm.

Reichweiten

Fehlende bzw. umstrittene Meßverfahren führen dazu, daß in der Literatur, insbesondere in Fachzeitschriften, stellenweise wahllos mit Begriffen zu Zugriffszahlen hantiert wird. Daher sollen hier nur solche Online-Angebote von Rundfunkanstalten berücksichtigt werden, die Nutzerzugriffe in der Einheit Page Views beziffern – u.a. auch weil dieser Wert, der derzeit am häufigsten ausgewiesene Wert ist. Die hier im folgenden aufgeführten Angaben zu Nutzerzugriffen basieren – mit Ausnahme von RTL und PRO SIEBEN – sämtlich auf der Befragung 3–6/97.

Die Einheit Page Views bezeichnet die Anzahl der Sichtkontakte beliebiger Nutzer mit einer HTML-Seite. Die Summe aller Page Views liefert ein Maß für die Attraktivität des Gesamtangebots. Auf eine Ausweisung von Hits soll hier gänzlich verzichtet werden.

Aufgrund einer fehlenden „Online-Währung“ verzichteten viele Rundfunkanstalten – z.B. die ARD („Das Erste“), Radio FRITZ, der HR, RB und der WDR – ganz bewußt auf eine Ausweisung ihrer Zugriffszahlen in der Befragung 3–6/97. Der BR und der ORF machten lediglich Angaben zu Hits auf ihren Servern.

6 Evaluierung der Online-Angebote von Rundfunkanstalten

Im folgenden wird eine Evaluierung der Online-Angebote der Rundfunkanstalten aus dem RFA-Panel hinsichtlich der Umsetzung der als wesentlich erachteten Erfolgsfaktoren für Online-Marketing vorgenommen werden. Eine Klassifizierung und Auswertung der Online-Angebote der Sender bezüglich den Online-Marketinginhalten Advertainment, Infotisement und Benefitting scheint hierbei nicht sinnvoll, da die Grenzen zwischen den einzelnen Inhalten zu stark verwischen.

6.1 Erfolgsfaktor Interaktivität

Zunächst soll die Feedback- und Dialogorientierung der Online-Angebote der Rundfunkanstalten ausgewertet werden. Als Gradmesser dieser Orientierung wurde die Präsenz der Komponenten „E-Mail“ und „Gästebuch“ in den Web-Sites der Anstalten des RFA-Panels überprüft. Gemäß dieser Erhebung ist in alle 41 Online-Angebote eine E-Mail-Funktion integriert. Dabei haben 33 Sender die erste Ebene in der Seitenhierarchie (Homepage/Leitseite) mit der E-Mail-Funktion ausgestattet. Bei sechs Sendern ist die Feedbackfunktion auf der ersten Hierarchieebene unterhalb der Homepage (zweite Ebene) untergebracht und nur bei zwei Sendern mindestens zwei Hierarchieebenen (dritte Ebene und tiefer) unter der Leitseite. Die Angabe der Seitenhierarchieebene soll hierbei als Gradmesser dafür dienen, welche Priorität die Sender einer interaktiven Kommunikation mit dem Publikum beimessen. Zur Komponente „Gästebuch“ ist festzustellen, daß 23 Sender aus dem RFA-Panel über diesen Feedbackkanal verfügen, 18 Anstalten haben kein Gästebuch.

Weiterhin gilt es den Erfolgsfaktor Interaktivität anhand der Komponenten „Chats“ und „Gewinnspiele“ zu evaluieren – u.a. auch, um die Unterhaltungsorientierung der Angebote herausstellen zu können. Hierbei ist festzustellen, daß 20 Sender Gewinnspiele und 13 Sender Chats anbieten. Zehn Anstalten bieten beide Entertainment-Komponenten, 18 Anstalten weder die eine noch die andere an.

ABB. 7: FEEDBACKKOMPONENTE E-MAIL IN ONLINE-ANGEBOTEN VON RUNDFUNKANSTALTEN DES RFA-PANELS

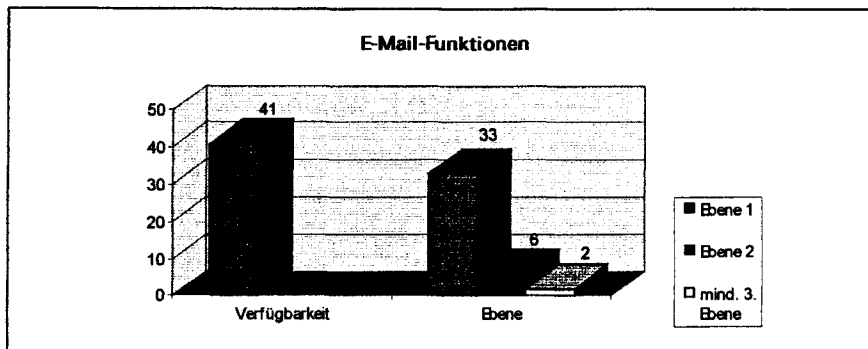
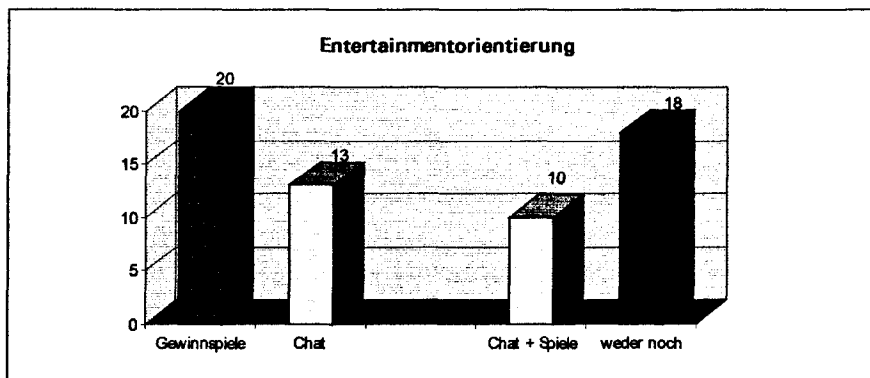


ABB. 8: ENTERTAINMENT-ORIENTIERUNG BEI ONLINE-ANGEBOTEN VON RUNDFUNKANSTALTEN DES RFA-PANELS

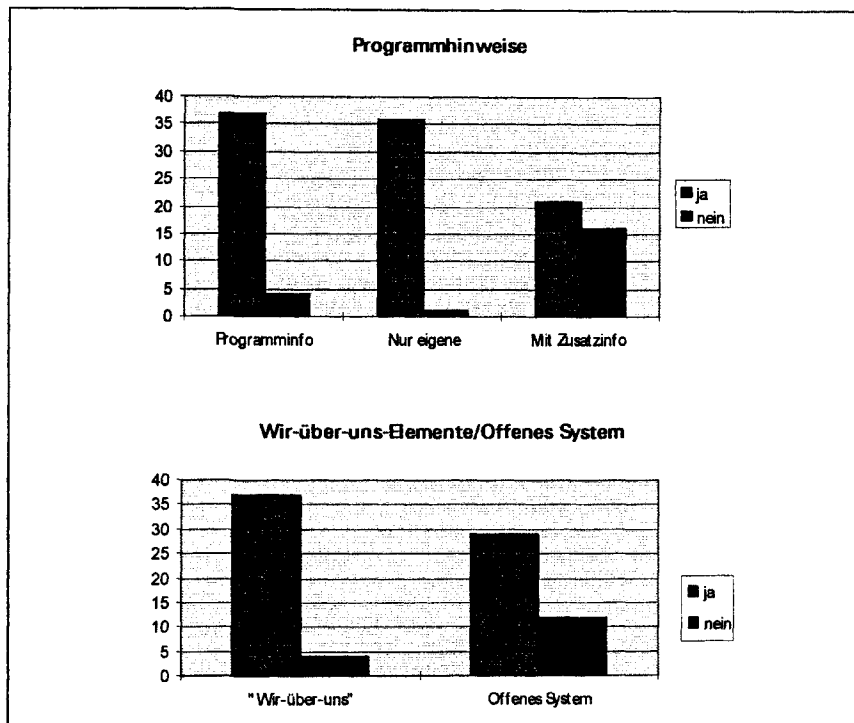


6.2 Erfolgsfaktor Mehrwerte

User akzeptieren Online-Angebote insbesondere dann, wenn dadurch ein Zusatznutzen – egal ob Unterhaltung oder Information – geboten wird. Als Indikatoren des Erfolgsfaktors „Mehrwert“ sollen die Komponenten „Programmhinweise“, „Hintergrundinformationen zu Programm, Machern und Organisation“ sowie „offenes/geschlossenes System“ herangezogen werden.

Weiterhin gilt es den Erfolgsfaktor Interaktivität anhand der Komponenten „Chats“ und „Gewinnspiele“ zu evaluieren – u.a. auch, um die Unterhaltungsorientierung der Angebote herausstellen zu können. Hierbei ist festzustellen, daß 20 Sender Gewinnspiele und 13 Sender Chats anbieten. Zehn Anstalten bieten beide Entertainment-Komponenten, 18 Anstalten weder die eine noch die andere an.

ABB. 9: MEHRWERTE DER ONLINE-ANGEBOTE VON RUNDFUNKKANSTALTEN DES RFA-PANELS



Weiterhin gilt es den Erfolgsfaktor Interaktivität anhand der Komponenten „Chats“ und „Gewinnspiele“ zu evaluieren – u.a. auch, um die Unterhaltungsorientierung der Angebote herausstellen zu können. Hierbei ist festzustellen, daß 20 Sender Gewinnspiele und 13 Sender Chats anbieten. Zehn Anstalten bieten beide Entertainment-Komponenten, 18 Anstalten weder die eine noch die andere an.

Unter anderem auch zur Überprüfung der Service-Orientierung und der Ausnutzung der Nichtlinearität des WWW wurde untersucht, ob die Web-Sites des RFA-Panels über externe Links zu anderen Web-Sites verfügen. Nicht einbezogen wurden hierbei allerdings kommerzielle Links bzw. Verweise, die den Anschein bezahlter Links hatten. Dabei war festzustellen, daß 29 Angebote „offen“, 12 dagegen „geschlossen“ sind.

Als nicht für Endkonsumenten (bzw. das Hörer- und Zuschauer-Publikum), sondern für die Business-to-business-Kommunikation identifizierbarer Mehrwert

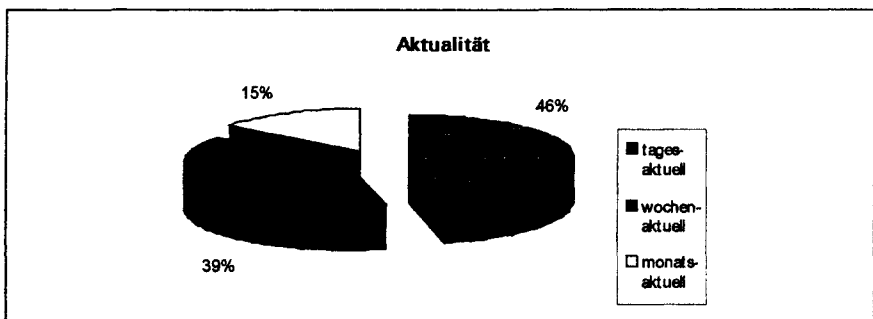
ist die Komponente „Informationen für Werbetreibende“ innerhalb des Online-Angebots ausweisbar. Über diese Option verfügen 18 Anstalten. Die Verfügbarkeit von Audio-on-demand- bzw. Video-on-demand- sowie RealAudio- bzw. -Video-Anwendungen im RFA-Panel wird in Kapitel fünf dokumentiert.

6.3 Sonstige Erfolgsfaktoren

Zu den sonstigen Erfolgsfaktoren, die hier evaluiert werden sollen, zählen „Aktualität“ und „Internationalität“. Die Bewertung der Aktualität der jeweiligen Online-Angebote wurde anhand der drei Kategorien „tagesaktuell“, „wochenaktuell“ und „monatsaktuell“ (Aktualisierung unregelmäßiger als wöchentlich) vorgenommen, wobei hier das jeweils letzte Updating – überprüfbar u.a. über „Letzte Meldungen“, Nachrichtenstand, Programmupdates etc. – als Indikator herangezogen wurde. Als Ergebnis dieser Erhebung sind 19 Sender als „tagesaktuell“, 16 als „wochenaktuell“ und sechs Sender als „monatsaktuell“ zu klassifizieren. Als wesentlicher Indikator des Faktors Internationalität eines Online-Angebotes kann das Vorhandensein einer Sprachoption herangezogen werden.

Hinsichtlich dieses Erfolgsfaktors ist anzumerken, daß nur fünf der 41 RFA-Panel-Anstalten über Sprachoptionen verfügen. Darunter der *WDR*, das *ZDF* und *3Sat* mit jeweils englischen Sites. Das grenzübergreifende *ARTE* gibt es in französisch sowie deutsch; ebenso präsentiert sich das Online-Angebot der *Deutschen Welle* in mehreren Sprachen.

ABB. 10: AKTUALITÄT DER ONLINE-ANGEBOTE VON RUNDFUNKANSTALTEN DES RFA-PANELS



Ausblick

Die Untersuchung hat gezeigt, daß die Rundfunkanstalten das WWW als „drittes Medium“ für die Verbreitung von Informationen und Unterhaltung erkannt haben und ihren WWW-Auftritt auch entsprechend gestalten. So ist zwar in der Regel eine gewisse Kopplung mit den Radio- und Fernsehprogramm erkennbar, allerdings ist das bereits in vielen Fällen nicht mehr das bestimmende Element für die Gestaltung der WWW-Präsentation. Allerdings existieren bereits WWW-Auftritte, deren Inhalte vom konventionellen Programm völlig losgelöst sind. Ein Beispiel hierfür ist die virtuelle Stadt Funcity (www.funcity.de) von Radio ffN. Hier wird sogar im Radioprogramm Werbung für die Nutzung dieses Dienstes gemacht.

Bezüglich der Verschmelzung von WWW- und Fernsehtechnologie sind noch keine signifikanten Fortschritte auf Anwendungsebene erkennbar, obwohl die notwendigen Basistechnologien bereits verfügbar sind. Allerdings bieten bereits eine Reihe Radiosender das Mithören des aktuellen Radioprogramms mit RealAudio oder ähnlichen Systemen in mittlerweile einigermaßen akzeptabler Qualität an.

Es ist daher davon auszugehen, daß sich die Rundfunksender im WWW dauerhaft und flächendeckend etabliert haben, aber die großen Innovationen noch bevorstehen.

Literatur

[Fantapié Altobelli/Hoffmann 1996]

Fantapié Altobelli, C./Hoffmann, S.: *Werbung im Internet*, Kommunikations-Kompendium, Band 6, Unterföhring 1996

[Hünerberg/Heise/Mann 1996]

Hünerberg R./Heise G./Mann A.: *Handbuch Online-Marketing. Wettbewerbsvorteile durch weltweite Datennetze*, Landsberg/L. 1996

[Siegle 1997]

Siegle, J. A.: *Online-Marketing von Rundfunkanstalten*, Wiesbaden 1997.

The “Virtual Museum”¹: New Perspectives For Museums to Present Objects and Information Using the Internet as a Knowledge Base and Communication System

Werner Schweibenz

University of Saarland, Germany
Mainzer Str. 129
D-66121 Saarbrücken
w.schweibenz@rz.uni-sb.de

Content

- 1 Introduction
- 2 Museums and the Mass Media
- 3 Museum Information and Information Technology
- 4 The “Virtual Museum” and the Opportunities it offers
- 5 The Internet as a Communication and Knowledge Base for Museums
- 6 Conclusion

Abstract

The relationship between museums and mass media as well as the possible impact of information technology on museums are described. The “virtual museum” is defined as a means to establish access, context, and outreach by using information technology. The Internet opens the “virtual museum” to an interactive dialog with virtual visitors and invites them to make a virtual museum experience that is related to a real museum experience. Some research is described on how the Internet can be used as a knowledge base and trends from surveys how museums and virtual visitors use the Internet as a communication tool.

¹ In this paper the term virtual museum is used in quotation marks because to my knowledge there exist no accepted definition and the term in itself is an oxymoron.

Zusammenfassung

Die Beziehung zwischen Museen und Massenmedien wird ebenso dargestellt wie mögliche Auswirkungen von Informationstechnik auf Museen. Das "virtuelle Museum" wird definiert als ein Mittel, das mit Hilfe von Informationstechnik Zugang, Kontext und Kontaktaufnahme zu Besuchern ermöglicht. Das Internet öffnet das "virtuelle Museum" für einen interaktiven Dialog mit virtuellen Besuchern und lädt sie ein, im "virtuellen Museum" Museumserfahrungen zu machen, die mit dem realen Museum in Beziehung stehen. Es wird berichtet, wie Museen und virtuelle Besucher das Internet als Knowledge Base und Kommunikationsmittel benützen.

1 Introduction

Internet has become a household word. Even institutions with more conservative attitudes towards information technology such as museums start to recognize the possibilities it offers for the dissemination of information to a world wide public. The interest in the application of interactivity and multimedia in the museum community was stated in the *International Conferences on Hypermedia and Interactivity in Museums (ICHIM)* which started in 1991. For a lot of museums the next step was to establish a presence on the World Wide Web and bring their collections on the Information Highway. More and more museums are interested in using the Internet and in sharing their experiences. This is shown in the annual conference *Museums and the Web* which was established in 1997. In this situation it is necessary to ask what new perspectives information technology can offer to museums for presenting objects and information and how the Internet can be used as a knowledge base and communication system.

This paper intends to show some perspectives from museology and information science literature and it also discusses the question stated above, referring especially to the art museum. This kind of museum is of special interest because in the art museum the object has a status as a work of art and curators often think it speaks for itself and no additional information is necessary whereas the public would appreciate background information in order to understand the museum objects. As a start it is necessary to take a look at the relationship between the museum and the mass media. Researchers have found out that there exist some interesting parallels.

2 Museums and Mass Media

Museums are institutions with a lot of facets depending on the topics they deal with, e.g. art, culture, history or science to name just a few. Therefore a definition for "museum" can only refer to the general aspects that all museums have in common. Burcaw (1975), for example, offers ten definitions collected

from different authors and museum organizations. According to Burcaw (1975, pp. 9f), a museum can be defined as a nonprofit institution that collects, preserves and displays objects for educational or aesthetic purposes. By definition, a museum is bound to physical objects. But at the same time it has the function of communication and dissemination of knowledge, a mission stated as early as 1846 when the Smithsonian Institution was founded "for the increase and diffusion of knowledge among men" (Waidacher 1993, p. 98).

Dealing with the communicative function of the museum, some researchers think that there are interesting parallels between museums and the mass media. Sharon MacDonald (1996, p. 5) emphasizes that "museums clearly have much in common with other institutions and media". According to Roger Silverstone (1988, p. 231), "museums are in the communication business" and offer a display of objects and artifacts which has been designed to educate, inform and entertain. Silverstone (1994, p. 162) further states that "museums are in many respects like other contemporary media. They entertain and inform; they tell stories and construct arguments; they aim to please and to educate; they define, consciously or unconsciously; effectively or ineffectively, an agenda; they translate the otherwise unfamiliar and inaccessible into the familiar and accessible." Relating these tasks of the museum to mass media culture of the late twentieth century Silverstone asks what kind of medium the museum is and how museums relate to and participate in electronically mediated culture (1994, pp. 161f). The problem in answering this question is, according to Silverstone, that there are obvious differences between museums and broadcast media such as newspapers, radio or television because museums occupy physical spaces, contain objects and encourage interactivity while they allow the visitor to wander through their exhibits. This range of interaction is not available to broadcast media that can only mediate between spaces but do not offer control (Silverstone 1988, p. 235).

Although there are differences in the degree of interaction between museums and broadcast media, there are striking parallels in behavior between museum audience and mass media audience as the research of the sociologist Heiner Treinen (1980; 1993) shows. Treinen has identified a kind of behavior that people visiting museums and people using mass media have in common. He calls it the phenomenon of "active dozing", a purposeless, planless activity that looks for gaining and maintaining permanent stimulation/diversion (1993, p. 89). In museums "active dozing" appears in the form of "cultural window shopping", which means that "visitors behave as if museums were mass-media: they linger in particular in front of objects with which they have already, before the museum visit, had something to do. As long as they know something about the category of objects and thus about the structure of the display, then a few glances may suffice to check off a few points, to stimulate thought, to supplement knowledge. Everything else is simply 'taken on board', or treated as mere diversion" (Treinen 1993, p. 90). This indicates that most museum visitors do not get the full value out of their visit and that museums could be more attractive to visitors if they would provide more information and entertainment or

a combination of both - edutainment. Especially the combination of information and entertainment is important because museums are in competition for visitors with other leisure pursuits and mass media (MacDonald, Sharon 1996, p.1) that use information technology extensively in order to become attractive to the audience. George MacDonald and Stephen Alsford (1995, pp. 129f) describe how museums compete with theme parks for recreational and leisure-time activities and how they adopt certain means usually applied by theme parks to attract visitors. Their conclusion is that there are new structures in communication and learning that create a dynamic interplay between the two and that both will pursue their special ways of presenting and interpreting information (1995, pp. 143-145).

3 Museum Information and Information Technology

In the 1980s a shift of paradigms happened in museology: the importance of objects was questioned in favor of the importance of information (Pearce 1986). Scholars like Wilcomb E. Washburn suggested that the emphasis of museum work should be put on information rather than on objects (1984, pp. 14f), others like George MacDonald and Stephan Alsford (MacDonald/Alsford 1989, MacDonald/Alsford 1991, Alsford 1991, MacDonald 1992) described the museum as an information utility going as far as stating that museums need to think of information, rather than of material objects, as their basic resource (Alsford 1991, p.8). Finally museums were no longer thought of as being repositories of objects only but as "storehouses of knowledge as well as storehouses of objects" (Cannon-Brookes 1992, p. 501; Hooper-Greenhill 1992, pp. 3f).

An important issue related to that shift of paradigms was the growing importance of museum education and visitor studies. A number of studies on visitor interests were conducted like those of Melora McDermott (1988) and the Getty Center for Education and the Arts (1991). The findings of these reports showed that the visitors thought information to be important for the appreciation of museum objects in general and for art in particular. If they do not get this information, visitors lack the key to the understanding of museum objects and they cannot connect to the object. Therefore they behold the objects for a few seconds only (Treinen 1996, p. 65). As Laura Chapman states (1982, p. 48), the myth that "objects speak for themselves" forgets that the meaning of an object is learned and established by the context. The importance of context and how it is communicated is also emphasized by Treinen (1996, p. 65) who reasons that communication is the key to the understanding of the museum object. Instead of only presenting objects, museums have to create meaning and establish context. This shows that an important aspect of the museum is to connect visitors, objects and information, an idea for which Glen H. Hoptman has introduced the term connectedness.

According to Hoptman (1992) connectedness is the basic feature of the "virtual museum", as it seeks to describe the interrelated and interdisciplinary presentation of museum information with the help of integrated media. Connectedness is the quality that allows the "virtual museum" to transcend the abilities of the traditional museum in presenting information. This new quality can be realized in different ways, for example, in displaying digital representations of works of art next to comparative works by the same artist, artists who have influenced him or her, or works of the same style or period that are exhibited in museums at various geographic locations or that are otherwise not normally accessible together. Hoptman's concept of connectedness coupled with the following quote shows the value of the "virtual museum":

The concept of the Virtual Museum demonstrates how limitations imposed by the traditional method of organizing and presenting information can be overcome in the context of museum visits. In a nutshell, the Virtual Museum provides multiple levels, perspectives, and dimensions of information about a particular topic: it provides not only multimedia (print, visual images through photographs, illustrations or video, and audio), but, more important, it provides information that has not been filtered out through these traditional methods (Hoptman 1992, p. 146)

Connectedness does not merely mean to link objects together but to give visitors the opportunity to focus on their special interests by pursuing them in an interactive dialog with the museum. This is an important step in the development from the traditional museum to the museum of the future, as Hooper-Greenhill (1994, pp. 134) emphasizes: the museum changes from a "collection-driven museum" to an "audience-driven museum" that tries to relate to particular visitor groups and to focus on the visitors instead of the collections they visit. Ben Davis (1994, p. 70) reaches the same conclusion when he states: "The digital museum can be visitor-centered rather than curator-centered."

An important step towards the "audience-driven" or "visitor-centered" museum is that museums try to reach out to their prospective visitors. The possibilities of outreach that telecommunication services offer for the museum to link with its visitors are emphasized by several researchers (Anderson 1997; Argoski 1995; Bearman 1995a, 1995b; Bowen/Bennett/Johnson 1998; MacDonald/Alsford 1997). For them, the Internet seems to be the ideal knowledge base and communication system to achieve this goal. The World Wide Web offers the possibility to link text, images, sound and video to an interactive hypermedia setting which promises interesting opportunities for museums to present objects and information and offer remote access to their collections. This will have a deep impact on the traditional museum. Both Maxwell L. Anderson (1997, p. 27) and George MacDonald and Stephan Alsford (1997, pp. 267f) think that the museum will combine its role as a repository with the use of telecommunication technologies adding a new, digital dimension to the traditional museum. This

digital dimension will lead to a new form of museum that enriches the objects with information: the "virtual museum".

4 The "Virtual Museum" and the Opportunities it offers

The idea of the "virtual museum" is currently under construction. In the museum and information science literature a variety of terms are used synonymously for museum-related digitized information resources, e.g. electronic museum, digital museum, on-line museum, hypermedia museum, meta-museum, Web museum, and Cyberspace museum. All these terms share the concept of digitized museum information that is brought together in an online-accessible collection.

A well-known definition for the "virtual museum" was presented by Geoffrey Lewis (1996, WWW) on the museum discussion list "museum-l" and in the Article Section of Britannica Online, the Internet version of the Encyclopaedia Britannica. Lewis described the "virtual museum" as

a collection of digitally recorded images, sound files, text documents, and other data of historical, scientific, or cultural interest that are accessed through electronic media. A virtual museum does not house actual objects and therefore lacks the permanence and unique qualities of a museum in the institutional definition of the term. (Britannica Online, Article Section, 1996)

Lewis emphasized the important difference between the traditional museum and the "virtual museum", i.e. the real and the digital object. Related to this is the lack of unique qualities to which Walter Benjamin refers to as "aura" in his famous 1936 essay "The Work of Art in the Age of Mechanical Reproduction". As the research of M. Travis DiNicola (1995, WWW) shows Benjamin's essay is relevant for the digital reproduction of works of art. The heated discussion on "aura" in the museum literature gives proof of a dispute between progressive researchers who grant digitally reproduced objects a kind of auratic quality (DiNicola 1995, WWW; Douglas Davis 1995, p. 381-385) and traditionalists (Schäfer 1995, pp. 75f; Mitchell/Strimpel 1997, p. 32) who refuse to grant it.

This aspect may be important for the museologists and specialists but not so much for the common museum visitors. The research of John Falk and Lynn Dierking (1992 p. 2f) shows that from the visitor's perspective the museum experience consists of three contexts: the personal context which incorporates a wide range of personal experience, knowledge and motivation; the social context which refers to the social environment in which the visit happens; and the physical context which relates to the architecture of the building as well as to the objects contained within. As Falk and Dierking (1992, p. 13f) emphasize the "decision to visit a museum involves matching personal and social interests and desires with the anticipated physical context and the associated activities of a museum" and that visitors want to see content in context. But as stated earlier in regard to Treinen's theory of "cultural window shopping", most museum

visitors do not get the full value out of their visit. In this context, the research of Falk and Dierking (1992, p. 37) offers an interesting insight: the visitor's museum experience highly depends on expectations of the visitor and how closely they fit to the visitor's museum agenda and the actual museum experience. If the visitor has "informed expectations" about the museum visit, there will be a close fit between the expectations and the actual visit. Eventually this will lead to a positive, reinforcing attitude about museums.

All these arguments lead to the question that is crucial to the "virtual museum": "Any discussion of virtual museums must begin with the question: Is it possible for the patron to have a meaningful or 'real' experience visiting a 'virtual' museum?" (Argoski 1995, WWW). This question is partly answered by Falk and Dierking (1998) and Lynne Teather (1998). Falk and Dierking (1998, p. 8f) draw parallels between museum-going and visiting museum Web sites and suggest – based on the modest research that is available – that creating a Web experience is as complex a behavior for virtual visitors as museum going because both are centered on free choice learning. Teather refers to the "user-created experience of museums", as she calls it and states (Teather 1998, p. 6) that "the essence of the museum experience that we wish to transfer to the web [...] is about meaning and knowledge building that is based in the visitor". According to Teather, the museum experience – both the real and the virtual one – is centered around meaning making of the visitors, something for which they need information. But while Teather values information, she doesn't support "a philosophy of information for information's sake" (p. 8f) but a philosophy of concepts. The idea of concepts is especially important because access to the information is not enough as Kevin Donovan (1997, p. 128) points out. He argues that "access to much of our on-line sources is of little value because museums add so little value to the data they provide". The object data with no value added are not useful for those who can not interpret and analyze it. Therefore Donovan (1997, p. 130) suggests that museums should give up their object-centric manner in the electronic space because the object is only available in a surrogate form. Instead they should present stories of the culture, historical context, people and places the object is related with. This leads back to Hoptman's concept of connectedness.

Taking all these ideas into account the "virtual museum" can be defined as follows: The "virtual museum" is a logically related collection of digital objects composed in a variety of media, and, because of its capacity to provide connectedness and various points of access, it lends itself to transcending traditional methods of communicating and interacting with the visitors being flexible toward their needs and interests; it has no real place or space, its objects and the related information can be disseminated all over the world.

This definition follows the research of James Andrews and Werner Schweibenz (1998) and describes the "virtual museum" essentially as a museum without walls, as Bearman (1992, p. 126) calls it. This "virtual museum" opens itself to an interactive dialog with visitors offering them connected digital objects and

information that is readily accessible from outside the museum. At its best, the "virtual museum" connects the visitors with valuable information across the entire globe and gives them a dynamic, multidisciplinary and multimedia approach to the collection as Jamie McKenzie puts it (McKenzie 1997, WWW). The idea is not a new one, it was already described in a scenario presented by Allon Schoener (1968, p. 364) at the 1968 *Conference on Computers and their Potential Application in Museums*. Now the Internet offers the possibility to realize it.

5 The Internet as a Communication and Knowledge Base for Museums

The prerequisite for the "virtual museum" are digitized data. Therefore it is important that museums "recognize that the methods of information processing and delivery in the next century will be digital, and they should begin now to build digital image and knowledge bases reflecting their holdings and the contexts of the creation, discovery, use, and meaning" (Bearman 1992, p. 135). This is necessary because the collection of the "virtual museum" consists of those objects only that are documented in the museum's databases and not of those held in its storerooms, therefore undocumented or inadequately documented objects are lost to the virtual collection as Bearman (1995a, p. 21) states. In the long run, Bearman argues, the difference between having objects "on exhibit" and "on display" will diminish and the whole collection will be accessible online all the time. The goal will be, according to Argoski (1995, WWW), that the information collected by the museum will "be reused in a variety of ways and through different media". Argoski calls this "repurposability" and thinks it of prime importance to the "virtual museum".

The Fine Arts Museums of San Francisco (FAMSF) offer an interesting perspective of how digitized museum information can be used in different contexts. A huge part of the museums' holdings are digitized in the imagebase "The Thinker" (URL = <http://www.thinker.org/imagebase/index.html>) which contains over 70,000 images and related information (for details see Schweibenz 1998). The museum staff uses the imagebase to plan exhibitions, researchers who visit the museum use it to select works of art they want get from storage, visitors access information on objects while visiting an exhibit, and virtual visitors visit the imagebase on the Web. The next step is to combine the collection in "The Thinker" with that of other museums in the digital library of the *Art Museum Image Consortium* of which the FAMSF is a member.

The *Art Museum Image Consortium* (AMICO, URL = <http://www.amn.org/-AMICO/>) is a consortium of 23 North American art museums that was founded in October 1997. The members plan to set up an Art Museum Image Consortium Library that shall contain the documentation of over 20,000 works of art by the end of the year 1998 and be twice the size in 1999. The AMICO Library consists of multimedia documentation of works of art, including digital images, collection data, curatorial records, scholarly research and educational

material. In general the images will be high resolution pictures (1024 x 768 pixels in 24 bit color), some images will be available in resolution up to twenty times the minimum, allowing a zooming function. Sample records are available on the AMICO Web site. The AMICO Library will be licensed to educational institutions (K-12, universities, museums, and public libraries) and the revenue is shared by the members and used to meet the AMICO technical specifications.

The AMICO Library can be regarded as an example for a "virtual museum" that invites visitors to make a virtual museum experience. It fulfills the predictions of Howard Besser who speculated in his 1987 article *The Changing Museum* (Besser 1987, p. 14) that the role of the "museum will change from a static repository of information (akin to an archive) to a more dynamic, interactive information source (more like a library)", changing public perception of the objects they store. Museums will become more accessible and more democratic as a wider public gets involved (Besser 1987, p. 16f). Besser argues that the museum is a tightly controlled environment has limited hours, admission fees, and an elitist way of looking at art. But things change with the advent of information technology which allows increased access and interactivity. This will shift the museum from being a passive repository to a more active role. As Roland Jackson et al. (1998, p. 1) point out, "the capacity of the Internet, synchronous and asynchronous, makes it possible to make contact and develop long-term relationships with the public". Another example is the Virtual Reference Desk (http://www.nmaa.si.edu/ref_desk.html) of the National Museum of American Art (NMAA) where the reference librarian answers questions from researchers and laymen from all over the U.S. The virtual reference desk was intended as a facet of interactive features which included online chats, message boards, entertaining contents, suggestion boxes and the digital reference service. These features were thought as a means to interact with the audience in order to find out what museum services they most wanted, as Joan Stahl, the reference librarian of the NMAA, puts it (Stahl 1998, p. 11). The electric reference service is well accepted as the average of approximately 350 questions per month shows, the users range from students to artists and professionals both inside and outside the art fields (p. 11f).

Other ways of using the Internet as a means of communication are described by Jonathan P. Bowen, Jim Bennett, Jim and James Johnson (1998, WWW) who did some research in how museums have been using the Internet and how they could use it in the future. They sum up several reasons why museums should use the Internet:

- Maintaining a presence on the Internet provides the potential for worldwide publicity.
- The Internet offers fast and convenient communication with both colleagues and the public.
- Virtual exhibitions can mirror and keep up actual exhibitions in the galleries. This is what around three quarters of virtual visitors expect to find at

museum Web sites. Moreover virtual exhibitions allow access to material that is otherwise not available.

- The Internet offers an alternative, cheap, and complementary form of information provision and is likely to act as a draw for prospective visitors who would like to see the real thing.
- The Internet offers remote access to scholarly research of collections in online databases.

The research of Bowen, Bennett and Johnson shows that the Internet can be a useful tool for communication with the audience. But there is only little research available both on how museums and virtual visitors use the World Wide Web.

Stephanie James (1997, WWW) has done a survey on how museums use their Web site which included 33 museums around the world. For 30 percent the rationale behind the creation of the Web site was promotion/marketing, for 19 percent to have a presence on the Web, and for 11 percent it was educational reasons. Over times the purpose has evolved with 73 percent of the participating museums and shifted away from the promotion/marketing emphasis. Now for 38 percent the emphasis is now greater access to collections, for 23 percent greater educational focus, for 15 percent increasing interactivity. At the same time 54 percent of the participating do no web-visitor research at all and 14 percent do hit counting. Only 23 percent investigate the length of visit and the pages looked at. This shows that there is still a lot of research to do for museum and information science professionals.

While visitor studies are an emerging discipline in the Anglo-American museum world, research on virtual visitors is still a comparatively new field. Katherine Futers (1997, WWW) presents the data of a survey conducted by Rachel Reynolds, a museum studies student at the Leicester University. According to this survey 88 percent of the virtual visitors are based in North America, women make up 46 percent of them and the average age of museum web page visitors is 40 to 64 years of age. 74 percent of the virtual visitors expect to find online exhibition on museum Web pages, 87 percent expect images and 52 percent want to download these images – a nightmare for most museums' copyright concerns. Currently a survey of a student of the University College London is going on (International Internet Museum Survey Questionnaire <http://geocities.com/SoHo/Museum/8355/>). How virtual visitors use the World Wide Web and what they expect to find there will hopefully soon become an evolving field for visitor studies.

These survey data can only indicate trends but they give hints how the World Wide Web can evolve to a communication system between museums and visitors. These data are supported by the hit counts of Web catalogs that offer access to museum Web sites. An example is the site *Museums around the world* of the International Council of Museums (<http://www.icom.org/vimp/world.html>) respectively the *Virtual Library of museums page* (<http://www.comlab.ox.ac.uk/-archive/other/museums/>) both maintained by Jonathan P. Bowen. Bowen has counted more than 2.5 million visitors since

the sites were established. Bowen's statistics (Bowen 1995, p. 38f) show that museum information on the Web is on the par with American Football. This means that museums on the Web are doing reasonably well as far as the numbers of visitors are concerned.

6 Conclusion

The use of telecommunication technologies offers interesting perspectives for museums and the opportunity to add a new, digital dimension to the traditional museum, thereby creating a "virtual museum". The foundations for the "virtual museum" are already laid. Bearman (1995b, pp. 15f) estimates that by the end of this decade over 20 million original objects will have been digitized. In this way, museums and the digital information they offer will become loadstones of content for the growing multimedia industry and for museum initiatives for outreach to the public (Bearman 1995a, p. 12). As some statistics and research suggest, the public looks for and appreciates museum information on the Internet but has high standards that the museums have to meet. The Internet is a great opportunity which the museums should use to broaden its audience. So there is a piece of advice for museums concerning the Web: "Be there or be square!"

References

[Alsford 1991]

Alsford, Stephen. *Museums as Hypermedia: Interactivity on a museum-wide scale*. Hypermedia & interactivity in museums: proceedings of an international conference. Ed. David Bearman. Pittsburgh, PA: Archives and Museums Informatics, 1991. 7-16

[Art Museum Image Consortium undated]

Art Museum Image Consortium (AMICO). Internet, URL = <http://www.amn.org/-AMICO/>. Version: undated. Last visited: 08/30/1998

[Anderson 1997]

Anderson, Maxwell L. *Introduction*. The Wired Museum- Emerging Technology and Changing Paradigms. Ed. Katherine Jones-Garnil. Washington, D.C.: American Association of Museums, 1997. 11-34

[Andrews/Schweibenz 1998]

Andrews, James/Schweibenz, Werner. *A New Media for Old Masters: The Kress Study Collection Virtual Museum Project*. Art Documentation 17(1), 1998. 19-27

[Argoski 1995]

Argoski, Jason. *Virtual Museums: The Web Experience*. The Virtual Mirror. Internet, URL = <http://www.vmirror.com/rov-int/museums.html>. Version: 12/17/1995. Last visited: 07/29/1998

[Bearman 1995a]

Bearman, David. *Information Strategies and Structures For Electronic Museums*. Ed. Museum Documentation Association (Great Britain): Information: The Hidden Resource, Museums and the Internet. 7th

Conference 1995 Edinburgh, Scotland. Cambridge, England: Museum Documentation Association, 1995. 5-22

[Bearman 1995b]

Bearman, David. *Museum Strategies for Success on the Internet*. Ed. Giskin Day: Museum Collections and the Information Highway. Proceedings of a Conference on Museums and the Internet 10 May 1995. London: Science Museum, 1995. 15-27

[Bearman 1992]

Bearman, David. *Interactive Multimedia in Museums*. Eds. Susan Stone and Michael Buckland, Michael: Studies in Multimedia: State-of-the-Art Solutions in Multimedia and Hypermedia. Proceedings of the 1991 Mid-Year Meeting of the American Society for Information Science San Jose, California, April 1991. Medford, NJ: Learned Information, 1992. 121-137

[Benjamin 1985]

Benjamin, Walter. *The Work of Art in the Age of Mechanical Reproduction*. Illuminations. Walter Benjamin - Essays and Reflections. Ed. Hannah Arendt. New York, NY: Schocken Books, 1985. 217-51

[Besser 1987]

Besser, Howard. *The Changing Museum*. Information: The Transformation of Society. ASIS'87 Proceedings of the 50th Annual Meeting of the American Society for Information Science, Boston, MA, October 4-8, 1987. Ed. Chen, Ching-chih. Vol. 24. Medford, NJ: Learned Information, 1987. 14-19

[Bowen 1995]

Bowen, Jonathan. *The Virtual Library of Museums*. Ed. Giskin Day: Museum Collections and the Information Highway. Proceedings of a Conference on Museums and the Internet 10 May 1995. London: Science Museum, 37-39.

[Bowen/Bennett/Johnson 1998]

Bowen, Jonathan P./Bennett, Jim/Johnson, James. *Virtual Visits to Virtual Museums*. Internet, URL = <http://www.museums.reading.ac.uk/mw98/paper/>. Version: April 1998. Last visited: 07/28/1998

[Britannica Online 1998]

Britannica Online, Articles of the Year, 1996. Internet, URL = <http://www.eb.co.uk:195>. Version: undated. Last visited: 05/08/1998

[Burcaw 1975]

Burcaw, George Ellis. *Introduction to Museum Work*. Nashville, TN: American Association for State and Local History, 1975

[Cannon-Brookes 1992]

Cannon-Brookes, Peter. *The Nature of Museum Collections*. Manual of Curatorship, 2nd. ed. Ed. J. Thompson. London: Butterworth, 1992. 500-512

[Chapman 1982]

Chapman, Laura. *The Future and Museum Education*. Museum News, July/August 1982. 48-56

[Davis, Ben 1994]

Davis, Ben. *Digital Museums*. Aperture. No. 136 Summer, 1994. 68-70

[Davis, Douglas 1995]

Davis, Douglas. *The Work of Art in the Age of Digital Reproduction*. Leonardo (USA), 28(5), 1995. 381-386

[DiNicola 1995]

DiNicola, M. Travis. *The Work of Walter Benjamin in the Age of Digital Reproduction*. Internet, URL = <http://cac.psu.edu/~mtd120/palmer/thesis/benjamin.html>. Version: 11/04/1995. Last visited: 07/27/1998

[Donovan 1997]

Donovan, Kevin. *The Best of Intentions: Public Access, the Web and the Evolution of Museum Automation*. Ed. David Bearman and Jennifer Trant: *Museums and the Web 1997: Selected Papers*. Proceedings of the First International Conference Los Angeles, California, March 16-19, 1997. Pittsburgh, PA: Archives & Museum Informatics, 1997. 127-133

[Falk/Dierking 1992]

Falk, John H./Dierking, Lynn D. *The Museum Experience*. Washington: Whalesback Books.

[Falk/Dierking 1998]

Falk, John H./Dierking, Lynn D. *Understanding Free-Choice Learning: A Review of the Research and its Application to Museum Web Sites*. Ed. David Bearman and Jennifer Trant: *Museums and the Web 1998: Selected Papers* published on CD-ROM. Proceedings of the Second International Conference 1998 Toronto, Canada, April 22-25, 1998. Pittsburgh, PA: Archives & Museum Informatics, 1998. Pages counted as printed.

[Futers 1997]

Katherine Futers. Tell Me What You Want, What You Really, Really Want: A Look at Internet User Needs. Proceedings of EVA'97 Paris, France. Internet, URL = http://www.open.gov.uk/mdocassn/eva_kf.htm. Version: 1997. Last visited: 08/30/1998

[Getty 1991]

Getty Center for Education and the Arts. *Insights, Museums, Visitors, Attitudes, Expectations, A Focus Group Experiment*. Los Angeles, CA: The Paul Getty Trust, 1991

[Hooper-Greenhill 1992]

Hooper-Greenhill, Eilean. *Museums and the Shaping of Knowledge. The Heritage: Care-Preservation-Management*. London: Routledge, 1992

[Hooper-Greenhill 1994]

Hooper-Greenhill, Eilean. *Museum education: past, present and future. Towards the Museum of the Future. New European Perspectives*. Ed. Roger Miles and Lauro Zavala. London/New York: Routledge, 1994. 133-146

[Hoptman 1992]

Hoptman, Glen H. *The Virtual Museum and Related Epistemological Concerns*. Ed. Edward Barrett: *Sociomedia. Multimedia, Hypermedia and the Social Construction of Knowledge*. Cambridge, Mass.: MIT-Press, 1992. 141-159

[Jackson et al. 1998]

Jackson, Roland/Bazley, Martin/Patten, Dave/King, Martin. *Using the Web to Change the Relation Between a Museum and its Users*. Ed. David Bearman and Jennifer Trant: *Museums and the Web 1998: Selected Papers* published on CD-ROM. Proceedings of the Second International Conference Toronto,

Canada, April 22-25, 1998. Pittsburgh, PA: Archives & Museum Informatics, 1998. Pages counted as printed.

[James 1997]

James, Stephanie. Museum Web Page Survey Results. Internet, URL = <http://www.chass.utoronto.ca/~sjames/museum/survey.htm>. Version: 03/05/1997. Last visited: 08/18/1998

[Lewis 1996]

Lewis, Geoffrey. *The Response of Museums to the Web*. Archives of MUSEUM-L@HOME.EASE.LSOFT.COM Museum discussion list. Internet, URL = <http://home.dc.lsoft.com/archives/museum-l.html>. Posted: 10/17/1996. Last visited: 07/01/1998

[MacDonald, George 1992]

MacDonald, George. *Change and Challenge: Museums in the Information Society*. Ed. I. Karp: Museums and Communications - The Politics of Public Culture. Washington, D.C.: Smithsonian Institution Press, 1992. 158-181

[MacDonald, Sharon 1996]

MacDonald, Sharon. *Theorizing Museums: An Introduction*. Ed. Sharon MacDonald, Sharon and Gordon Fyfe: *Theorizing Museums*. Oxford: Blackwell. 1-18

[MacDonald/Alsford 1997]

MacDonald, George/Alsford, Stephen. *Conclusion: Toward the Meta-Museum*. Ed. Katherine Jones-Garmil: *The Wired Museum - Emerging Technology and Changing Paradigms*. Washington, D.C.: American Association of Museums, 1997. 267-278

[MacDonald/Alsford 1995]

MacDonald, George/Alsford Stephen. *Museums and Theme Parks: Worlds in Collision?* International Journal on Museum Management and Curatorship, Vol. 14 No. 2, 1995. 129-147

[MacDonald/Alsford 1991]

MacDonald, George/Alsford, Stephen. *The Museum as Information Utility*. Museum Management and Curatorship, 10/1991. 305-311

[MacDonald/Alsford 1989]

MacDonald, George/Alsford, Stephen. *A Museum For The Global Village. The Canadian Museum of Civilization*. Hull: Canadian Museum of Civilization, 1989. See Chapters 6 The Museum as Communicator, 10 The Museum as Resource.

[McDermott 1988]

McDermott, Melora. *Through Their Eyes: What Novices Value in Art Experiences*. Annual Meeting Sourcebook. Washington, DC: American Association of Museums, 1988. 135-162

[McKenzie 1997]

McKenzie, Jamie. *Building a Virtual Museums Community*. Paper presented at the Museums & The Web Conference March 16-19, 1997 Los Angeles, California. Internet, URL = <http://fromnowon.org/museum/museweb.html>. Version: March 1997. Last visited: 08/28/1998

[Mitchell/Stripel 1997]

Mitchell, William J./Stripel, Oliver B. R. *To Be There Or Not To Be*

There: Presence, Telepresence and the Future of Museums. Museum News. Mar/Apr. 1997. 31-32 and 58-59

[Pearce 1986]

Pearce, Susan M. *Thinking about Things. Approaches to the Study of Artefacts.* Museum Journal, March 1986. 198-201

[Schäfer 1995]

Schäfer, Hermann. *Use and Embedding of Interactive Media in a Museum of Contemporary History.* Ed. David Bearman: Hands On Hypermedia and Interactivity in Museums. Selected Papers from the Third International Conference on Hypermedia and Interactivity in Museums (ICHIM '95 - MCN '95) San Diego, California October 9-13, 1995. Pittsburgh, PA.: Archives & Museum Informatics, 1995. 70-85

[Schoener 1968]

Schoener, Allon. *The Electronic Museum and Information Distribution.* Ed. Metropolitan Museum of Art: Computers and Their Potential Applications in Museums. A Conference Sponsored by the Metropolitan Museum of Art, April 15, 16, 17, 1968. New York, NY: Arno Press, 1968. 359 – 366

[Schweibenz 1998, forthcoming]

Schweibenz, Werner. *Museumsinformation im Internet am Beispiel der Webseiten zweier Kunstmuseen in den USA.* Ed. Marlies Ockenfeld. Proceedings des 50. Deutschen Dokumentartags. Frankfurt/M.: DGD, 1998.

[Silverstone 1994]

Silverstone, Roger. *The Medium is the Museum.* Ed. Roger Miles and Lauro Zavala: Towards the Museum of the Future. New European Perspectives. London/New York: Routledge, 1994. 161-176

[Silverstone 1988]

Silverstone, Roger. *Museums and the Media: A Theoretical and Methodological Exploration.* International Journal of Museum Management and Curatorship, Vol. 7, No. 3, 1988. 231-241

[Stahl 1998]

Stahl, Joan. "Have a question? Click Here": *Electronic Reference at the National Museum of American Art.* Art Documentation 17(1) 1998. 10-12

[Teather 1998]

Teather, Lynne. *A Museum is a Museum is a Museum ... Or Is it?: Exploring Museology and the Web.* Ed. David Bearman and Jennifer Trant: Museums and the Web 1998: Selected Papers published on CD-ROM. Proceedings of the Second International Conference Toronto, Canada, April 22-25 1998. Pittsburgh, PA: Archives & Museum Informatics, 1998. Pages counted as printed.

[Treinen 1996]

Treinen, Heiner. *Ausstellungen und Kommunikationstheorie.* Ed. Haus der Geschichte: Museen und ihre Besucher – Herausforderungen in der Zukunft. (Reihe Museumsfragen) Berlin: Argon Verlag. 60-71

[Treinen 1993]

Treinen, Heiner. *What Does The Visitor Want From A Museum? Mass-media Aspects Of Museology.* Ed. Bicknell, Sandra/Farmelo, Graham: Museum Visitor Studies in the 90s. London: Science Museum. 86-93

[Treinen 1980]

Treinen, Heiner. *Das Museum als Massenmedium – Besucherstrukturen, Besucherinteressen und Museumsgestaltung*. Ed. CECA/ICOM: Museumsarchitektur für den Besucher, 1980. 13-32

[Waidacher 1993]

Waidacher, Friedrich. *Handbuch der Allgemeinen Museologie*. Mimundus Bd. 3. Wien: Böhlau, 1993

[Washburn 1984]

Washburn, Wilcomb E. *Collecting Information, Not Objects*. Museum News, 62 February, 1984. 5-15

Wissensaustausch auf der Grundlage von Referenzmodellen – Anforderungen und Möglichkeiten

Andreas Dietzsch, Werner Esswein

Technische Universität Dresden, Fakultät Wirtschaftswissenschaften
Lehrstuhl Wirtschaftsinformatik, insbes. Systementwicklung
Münchner Platz, 01062 Dresden
dietzsch@wise.wiwi.tu-dresden.de
esswein@wise.wiwi.tu-dresden.de

Inhalt

- 1 Einleitung**
- 2 Begriffsbestimmung**
 - 2.1 Modellbegriff
 - 2.2 Aspekte des Referenzbegriffs
 - 2.2.1 Wiederverwendungsgrundlage
 - 2.2.2 Empfehlungscharakter
 - 2.2.3 Glaubwürdigkeit
 - 2.3 Schlußfolgerungen
- 3 Begriffliches Wissen**
 - 3.1 Ausgewählte Eigenschaften
 - 3.1.1 *Unschärfe und Kontextabhängigkeit*
 - 3.1.2 *Begriffshierarchien*
 - 3.2 Schlußfolgerungen
- 4 Systementwicklung mit Referenzmodellen**
 - 4.1 Anwendung im Wasserfallmodell
 - 4.2 Anwendung beim Prototyping
 - 4.3 Anwendung im Spiralmodell
 - 4.4 Schlußfolgerungen
- 5 Zusammenfassung**

1 Einleitung

Wissen wird heute immer mehr als Produktivkraft erkannt. Dies und das explosionsartige Anwachsen des weltweiten Wissens in allen Bereichen – man geht von einer Verdopplung alle fünf Jahre aus – machen den Transfer von Wissen zu einem zentralen Problem (vgl. (Probst+ 97), S. 21).

Man wird nur dann in der Lage sein dieses Problem zu lösen, wenn Techniken entwickelt und angewendet werden, die eine formale Beschreibung und Darstellung von Wissen erlauben.

Eine Form einer solchen formalen Beschreibung stellen Referenzmodelle dar. Die vorliegende Arbeit konzentriert sich auf die Betrachtung der Möglichkeiten des Einsatzes von Referenzmodellen zur Bereitstellung und Weitergabe von Wissen im Rahmen der Entwicklung betrieblicher Informationssysteme.

Anhand verschiedener Vorgehensmodelle bei der Systementwicklung wird untersucht, wie Referenzmodelle eingesetzt werden können und welche Auswirkungen auf die Systementwicklung sich daraus ergeben.

2 Begriffsbestimmung

Referenzmodelle scheinen bereits aus der Bedeutung des Begriffes heraus einander widersprechende Ziele zu verfolgen. Zum einen wird häufig auf die Allgemeingültigkeit und Übertragbarkeit der Lösungsbeschreibung hingewiesen. Dabei wird das Ziel verfolgt, eine Ausgangsbasis für die Erstellung von Individualmodellen (vgl. (Becker+ 96), S. 25) zu liefern.

Zum anderen besteht parallel dazu der Anspruch, einen möglichst konkreten Bezug zu Problemen z. B. einer Branche (vgl. (Reiter 98), S. 14) oder der Konfiguration eines Softwaresystems (vgl. (Scheer 98b), S. 41f.) herzustellen. Im folgenden wird untersucht, ob und auf welche Weise beiden Ansprüchen genügt werden kann.

2.1 Modellbegriff

Unter einem Modell wird allgemein eine vereinfachte Darstellung wesentlicher Strukturen bzw. Funktionen verstanden. Bei der Modellierung betrieblicher Informationssysteme wird die betriebliche Realität als Objektsystem in einem Modellsystem abgebildet (vgl. [Ferstl+98], S. 117ff.).

Eine Diskussion wird lediglich darüber geführt, wie Modelle aufzufassen sind: als Konstruktionen aus für relevant erachteten Elementen, wie u. a. in (Schütte 97) vorgeschlagen, oder als Abbildungen von Ausschnitten der Realität.

Für die weiteren Betrachtungen wird deshalb unter einem Modell die Festlegung eines Gestaltungsrahmens für die formale Beschreibung der Realwelt verstanden (vgl. (Körmeier 97), S. 5).

2.2 Aspekte des Referenzbegriffs

Der aus dem Lateinischen stammende Begriff *Referenz* ist mit mehreren Bedeutungen belegt und beschreibt u. a. eine Beziehung, Empfehlung oder auch glaubwürdige Auskunft (siehe auch (Fremdwörterbuch 84), S. 646). Es wird deutlich, daß der Begriff *Referenz* viel Raum für Interpretation läßt.

Die Diskussion über Referenzmodelle (vgl. z. B. (Winter 97)) rührt zum Teil daher, daß nicht geklärt ist, in welcher Bedeutung der Referenzbegriff auf das jeweilige Modell anzuwenden ist. Genaue Aussagen darüber müssen deshalb Bestandteil des Referenzmodells selbst sein.

2.2.1 Wiederverwendungsgrundlage

Unter einer Beziehung ist im Zusammenhang mit dem Einsatz von Referenzmodellen das Wiederverwenden oder die Adaption eines bereits modellierten Sachverhaltes zu verstehen. Dieser Aspekt entspricht der Auffassung vom Referenzmodell als Ausgangspunkt bei der Erstellung von Individualmodellen und wird häufig als wesentliche Eigenschaft von Referenzmodellen dargestellt (vgl. dazu auch (Becker+ 96), S. 25ff. und (Scheer 98a), S. 61ff.). In diesem Sinne ist jedes Modell oder Modellteil, das in einem anderen Modell wiederverwendet wurde, als Referenzmodell bzw. Referenzbaustein zu betrachten.

2.2.2 Empfehlungscharakter

Der Empfehlungscharakter von Referenzmodellen geht über die im vorangegangenen Abschnitt beschriebene Bezugnahme auf bereits modellierte Sachverhalte hinaus. Eine breite Basis von Wissen bildet dabei den zentralen Aspekt dieser Eigenschaft. Die Erstellung des Referenzmodells kann dabei auf praktischen Erfahrungen (siehe auch (Reiter 98), S. 4ff.), im Sinne der Beschreibung von Best Practice Cases (vgl. (Scheer 98a), S. 61) oder auf theoretischen Überlegungen (vgl. (Bihl+ 97)) basieren.

Unabhängig von der Art der Erstellung wird dem Anwender des Referenzmodells für einen konkreten Problembereich eine auf Expertenwissen basierende Lösung angeboten und deren Anwendung empfohlen. Als Beispiel sei an dieser Stelle auf das von der International Standards Organization (ISO) entwickelte ISO-OSI-Referenzmodell (siehe auch (ISO 83)) verwiesen.

2.2.3 Glaubwürdigkeit

Neben der Darstellung von Wissen in Informationsmodellen spielt deren inhaltliche und formale Qualität eine entscheidende Rolle für die damit entwickelten Informationssysteme. Eine Möglichkeit diese abzusichern, besteht in der Modellierung unter der Berücksichtigung von Leitlinien, wie z. B. den Grundsätzen ordnungsmäßiger Modellierung (vgl. [Becker+95]).

Modelle, die zu bestimmten Problemen übliche Lösungen dokumentieren, wie „Common Practice“-Referenzmodelle (vgl. (Reiter 98), S. 4), leiten ihren Referenzanspruch aus dem erfolgreichen Einsatz in konkreten Projekten ab. Verbunden damit ist die Annahme, daß ihre Glaubwürdigkeit im Sinne der Modellqualität dadurch nachgewiesen ist.

Die Eigenschaft der Glaubwürdigkeit ist jedoch nicht auf Aussagen über die Qualität von Informationsmodellen beschränkt. Auch die von Software-Herstellern angebotenen Referenzmodelle ihrer Systeme, z. B. das Referenzmodell zum R/3-System von SAP, lassen sich als glaubwürdige Auskunft über das System interpretieren.

2.3 Schlußfolgerungen

Die Analyse der Bedeutungen des Begriffs Referenzmodell hat gezeigt, daß sowohl dem Anspruch einer modellhaften Abstraktion als auch der Schaffung einer Bezugsbasis durch entsprechende Modelle genügt werden kann.

Es wurde weiter deutlich, daß der Referenzcharakter eines Modells wesentlich durch den Prozeß der Erstellung bzw. die Art der Verwendung des Modells bestimmt wird. Da der Referenzbegriff qualitativ unterschiedliche Interpretationen zuläßt, ist es notwendig, dessen Bedeutung im oben beschriebenen Sinne bei der Beschreibung des Referenzmodells anzugeben.

3 Begriffliches Wissen

„Durch die Verknüpfung bestimmter Begriffe mit bestimmten sprachlichen Ausdrücken verbindet sich das begriffliche Wissen mit der Sprache als einem externen Symbolsystem, und es entsteht die Möglichkeit einer sprachlichen Wissensvermittlung.“ ((Schnotz 94) S. 29)

Ausgehend von der Definition eines Modells als Gestaltungsrahmen für die formale Beschreibung von Ausschnitten der Realität (siehe auch Abschnitt 0) ist zu untersuchen, ob Kriterien existieren, anhand derer sich die Eignung eines Modells zum Austausch von Wissen erkennen läßt.

Dazu ist ein Rückgriff auf Ergebnisse der Analyse der bei der Wissensvermittlung wirkenden Prinzipien möglich. Diese werden durch Arbeiten auf den Gebieten der Linguistik, Sprachpsychologie und kognitiven Psychologie behandelt.

3.1 Ausgewählte Eigenschaften

3.1.1 Unschärfe und Kontextabhängigkeit

Systematische, invariante Zusammenhänge oder auch gemeinsame Funktionen stellen die Basis für die Bildung von Begriffen dar (vgl. (Schnotz 94), S. 23f.).

Es wird jedoch weiter darauf hingewiesen, daß die Bildung von Begriffen stets durch Unschärfe gekennzeichnet ist, die mit wachsender Komplexität des zu erfassenden Systems zunimmt. In diesem Zusammenhang wurde festgestellt, daß für viele Alltagsbegriffe keine Definitionen in Form fester Merkmalssätze angegeben werden können (vgl. (Schnotz 94), S. 25). Vielmehr ist die Kategorisierung, d. h. die Zuordnung bestimmter Objekte zu einer Kategorie, vom Kontext abhängig, in dem sie betrachtet werden (vgl. (Anderson 96), S. 153ff.). Unter Berücksichtigung dieses Sachverhalts müssen Forderungen nach Eindeutigkeit, wie z. B. bei (Nonnenmacher 94), S. 142, als nicht realisierbar angesehen werden.

Untersuchungen zur Wahrnehmung von Sprache und visuellen Szenen haben gezeigt, daß sich die Wahrscheinlichkeit einer korrekten Identifikation durch das Hinzufügen von Kontextinformation steigern läßt (vgl. (Anderson 96), S. 58ff.).

3.1.2 Begriffshierarchien

Angaben bezüglich der bei Referenzmodellen zu wählenden Abstraktions-ebenen existieren, wenn überhaupt, nur in qualitativer Form, wie z. B. „adäquater Abstraktionsgrad“ (siehe (Becker+ 96), S. 25).

Beim Verstehen komplexen begrifflichen Wissens werden zur Komplexitätsbewältigung Begriffe auf unterschiedlichen Abstraktionsstufen zusammengefaßt und in Beziehung zueinander gesetzt. Dabei war festzustellen, daß Begriffe auf mittleren Abstraktionsstufen, den höchsten Grad an Orientierung ermöglichen (vgl. (Schnotz 94), S. 30ff.).

Wesentliche Eigenschaften solcher Begriffe sind dabei die Reichhaltigkeit des Begriffsinhaltes und die daraus folgende gute Abgrenzbarkeit gegenüber anderen Begriffen. In diesem Zusammenhang wird auch von einem hohen spezifischen Informationsgehalt gesprochen.

3.2 Schlußfolgerungen

Durch den Einsatz von Referenzmodellen ist die Weitergabe von Wissen möglich, das z. B. bei der Modellierung betrieblicher Informationssysteme nutzbar ist. Referenzmodelle fungieren dabei u. a. als Mittel zur Normung und Standardisierung dieses Wissens auf unterschiedlichen Abstraktionsstufen.

Um die Unschärfe und Kontextabhängigkeit begrifflichen Wissens zu berücksichtigen, müssen Referenzmodelle Informationen über den Kontext der bei der Modellierung verwendeten Begriffe enthalten. Dies kann z. B. durch das Aufstellen von Namenskonventionen bzw. das Verwenden standardisierter Begriffsbausteine im Rahmen normsprachlicher Ansätze realisiert werden (vgl. (Becker+ 96), S. 75 und (ORS 98), S. 31ff.).

Dabei ist zu beachten, daß das Wissen auf hohen Abstraktionsebenen die Interpretation von Wahrnehmungseinheiten auf niedrigeren Ebenen bestimmt. Die Wissensrepräsentation muß deshalb durch Begriffe auf einer Abstraktionsebene erfolgen, welche sich, bezogen auf das jeweilige Darstellungsproblem, von anderen Abstraktionsebenen durch einen höheren spezifischen Informationsgehalt unterscheidet (vgl. (Anderson 96), S. 59ff.). Eine Operationalisierung dieser Abstraktionsebenen ist jedoch nicht möglich.

4 Systementwicklung mit Referenzmodellen

Im folgenden wird an drei Paradigmen der Softwareentwicklung der Einfluß des Einsatzes von Referenzmodellen untersucht.

4.1 Anwendung im Wasserfallmodell

Die Systementwicklung nach dem Wasserfallmodell basiert (in seiner ursprünglichen Form) auf einer sequentiellen Abfolge der fünf Phasen Analyse, Entwurf, Implementierung, Test sowie Wartung (vgl. (Pressman 92), S. 24ff.).

Der Einsatz von Referenzmodellen bietet sich bei diesem Vorgehen in zwei wesentlichen Bereichen an: in der Analyse- und Entwurfsphase sowie bei der Implementierung. Doch in der Test- bzw. Wartungsphase ist ein sinnvoller Einsatz möglich.

Aus der Eigenschaft von Referenzmodellen als Träger von Wissen über z. B. branchentypische Probleme und deren Lösungen, kann dabei von einer Verkürzung der Phasen sowohl während der Analyse des Diskursbereichs als auch des Entwurfs des Informationssystems ausgegangen werden.

In der Phase der Implementierung besitzen Referenzmodelle im Rahmen des Customizing von Standardsoftwaresystemen eine besondere Bedeutung.

Neben der Anpassung des Systems, wie z. B. des R/3-Systems durch das entsprechende SAP-Referenzmodell (vgl. (Scheer 98b), S. 41f.), werden die am Softwaresystem vorgenommenen Veränderungen im Referenzmodell festgehalten, mit direkten Auswirkungen auf die Wartungsphase, die von der Existenz und Qualität solcher Dokumentationen abhängt.

In der Testphase kann das Referenzmodell weiter als Bezugspunkt für das Feststellen von Abweichungen des realisierten Systems gegenüber den spezifizierten Anforderungen Anwendung finden.

4.2 Anwendung beim Prototyping

Den Kernpunkt des Prototyping (vgl. (Pressman 92), S. 29f.) bildet das Erstellen eines Modells des zu entwickelnden Softwaresystems. Dabei bildet die Definition allgemeiner Ziele - nicht jedoch detaillierter Anforderungen - den Ausgangspunkt des Entwicklungsprozesses.

Die Systementwicklung erfolgt in einem Kreislauf mit folgenden Phasen:

1. Aufstellung bzw. Verfeinern der Anforderungen - diese Phase bildet den Einstiegspunkt in ein Prototyping-Projekt
2. "Schnellentwurf" - dabei wird auf die Aspekte des Systems fokussiert, die für den Nutzer sichtbar sind, z. B. Ein- und Ausgabeformate
3. Erstellung des Prototypen
4. Evaluierung des Prototypen durch den Kunden
5. Verbesserung des Prototypen - aus dieser Phase heraus kann ein erneuter Einstieg in Schritt 2 erfolgen
6. Produkterstellung - diese Phase markiert das Ende des Projektes

Wie bereits beim Wasserfallmodell, können Referenzmodelle in der ersten Phase die Grundlage für das Aufstellen bzw. Verfeinern der Anforderungen an das System bilden. Ist das Referenzmodell so gestaltet, daß es als Ausgangsprototyp dienen kann, wird der erste Durchlauf der Phasen eins bis vier überflüssig.

Eine wesentliche, qualitative Verbesserung des Prototyping-Prozesses ist in der zweiten Phase zu erwarten, da durch die Vormodellierung des Systems im Referenzmodell auch wesentliche Aspekte darstellbar sind, die im endgültigen System für den Nutzer von Bedeutung, jedoch nicht zwangsläufig sichtbar sind, wie z. B. die Abwicklung bestimmter Verfahren unter Aspekten der Datensicherheit.

Die möglichen Effekte in der letzten Phase entsprechen denen, die in der Phase der Implementierung des Wasserfallmodells beschrieben wurden.

4.3 Anwendung im Spiralmodell

Die wesentliche Erweiterung des Spiralmodells nach (Boehm 88) gegenüber dem Wasserfallmodell bzw. Prototyping besteht im Hinzufügen einer Risikoanalyse.

Im Spiralmodell werden bis zum endgültigen Produkt die folgenden vier Bereiche iterativ durchlaufen:

1. Planung - dient dem Aufstellen von Zielen, Alternativen und Bedingungen
2. Risikoanalyse - dient der Analyse der Alternativen sowie der Identifizierung und Beseitigung von Risiken
3. Entwicklung - Erstellung des Produktes ausgehend vom ersten Prototypen bis zum endgültigen System
4. Evaluation durch den Kunden

In den Phasen der Planung und Entwicklung, in denen das Spiralmodell Gemeinsamkeiten mit dem Wasserfallmodell bzw. Prototyping aufweist, sind durch den Einsatz von Referenzmodellen die in den beiden vorangegangenen Abschnitten beschriebenen Effekte zu erwarten.

Bei der Risikoanalyse können Referenzmodelle zum Aufzeigen von alternativen Problemlösungen sowie zur Feststellung von Abweichungen des realisierten Systems vom Referenzmodell im Rahmen der Identifizierung von Risiken eingesetzt werden. Weiter ist es möglich, Erfahrungswerte über Risiken selbst im Referenzmodell zu dokumentieren.

4.4 Schlußfolgerungen

Die vorgestellten Beispiele machen deutlich, daß der Einsatz von Referenzmodellen die Dauer von Softwareprojekten verringern kann.

Dabei besteht die Möglichkeit, die Durchführung bestimmter Phasen zu beschleunigen (z. B. Analyse- oder Entwurfsphase im Wasserfallmodell) oder auf den Durchlauf bestimmter Phasen (z. B. erster "Schnellentwurf" im Prototyping) völlig zu verzichten.

Einen weiteren wesentlichen Aspekt des Einsatzes von Referenzmodellen stellt die Möglichkeit dar, qualitative Verbesserungen der Ergebnisse einzelner Entwicklungsphasen durch die Verfügbarkeit formalisierten, problemspezifischen Wissens zu erreichen, wie dies für die Phase des "Schnellentwurfs" im Prototyping gezeigt wurde.

5 Zusammenfassung

Der Begriff *Referenz* kann, in Bezug auf die Eigenschaften von Modellen, mit qualitativ voneinander abweichenden Bedeutungen verstanden werden. Dabei wird der Referenzcharakter eines Modells wesentlich durch den Modellierungsprozeß aber ebenso durch die Art der Verwendung des Modells bestimmt.

Bei der Weitergabe von Wissen fungieren Referenzmodelle u. a. als Mittel zur Normung und Standardisierung. Dies macht das Aufstellen von Namenskonventionen bzw. die Verwendung standardisierter Begriffsbausteine zu einem notwendigen Bestandteil von Referenzmodellen.

Die Abstraktionsebene, auf der die Wissensrepräsentation im Referenzmodell erfolgt, ist stets unter Beachtung des darzustellenden Problems zu bestimmen.

An konkreten Beispielen wurde verdeutlicht, daß die Durchführung von Softwareprojekten durch den Einsatz von Referenzmodellen beschleunigt und qualitativ verbessert werden kann. Dies führt jedoch dazu, daß die Qualität von Referenzmodellen für den Prozeß der Systementwicklung und für das entwickelte System von Bedeutung ist.

Nicht betrachtet wurde, inwieweit die Ableitung von Individualmodellen aus einem Referenzmodell selbst als Entwicklungsprozeß anzusehen ist und welche Konsequenzen sich daraus auf eine auf die Verwendung von Referenzmodellen ausgerichtete Systementwicklung ergeben.

Literatur

(Anderson 96)

Anderson, J. R.: *Kognitive Psychologie: eine Einführung*. 2. Auflage, Spektrum, Akademischer Verlag, Heidelberg, 1996.

(Becker+ 95)

Becker, J.; Rosemann, M.; Schütte, R.: *Grundsätze ordnungsmäßiger Modellierung*. In: *Wirtschaftsinformatik*, 37 (1995) 5, S. 435 - 445.

(Becker+ 96)

Becker, J.; Schütte, R.: *Handelsinformationssysteme*. verlag moderne industrie, Landsberg/Lech, 1996.

(Bihr+ 97)

Bihr, H.; Seelos, H.-J.: *Entwicklung eines Referenzdatenmodells für Krankenhäuser*. In: *Wirtschaftsinformatik* 39 (1997) 4, S. 367 - 371.

(Boehm 88)

Boehm, B.: *A Spiral Model for Software Development and Enhancement*. In: *Computer* 21 (1988) 5, S. 61 - 72.

(Ferstl+ 98)

Ferstl, O. K.; Sinz, E. J.: *Grundlagen der Wirtschaftsinformatik Band 1*. 3. Auflage, Oldenbourg, München, 1998.

(ISO 83)

International Standardization Organization: *Open Systems Interconnection – Basic Reference Model*. IST7498, ohne Ort, 1983.

(Körmeier 97)

Körmeier, K.: *Zielorientierte Festlegung von Geschäftsprozessen*. Dissertation, Dresden, 1997.

(Nonnenmacher 94)

Nonnenmacher, M. G.: *Informationsmodellierung unter Nutzung von Referenzmodellen: die Nutzung von Referenzmodellen zur Implementierung industriebetrieblicher Informationssysteme*. Lang, Frankfurt/M et al., 1994.

(ORS 98)

ohne Autor: *Unterlagen zum Methodenseminar*. Open Reference Systems GmbH, Potsdam, Mai 1998.

(Pressman 92)

Pressman, R. S.: *Software engineering: a practitioner's approach*. 3. Auflage, McGraw-Hill, New York et al., 1992.

(Probst+ 97)

Probst, G.; Raub, S.; Romhardt, K.: *Wissen managen: Wie Unternehmen ihre wertvollste Ressource optimal nutzen*. Gabler, Wiesbaden, 1997.

(Reiter 98)

Reiter, C.: *Einsatz von Referenzmodellen*. IDS Prof. Scheer Gesellschaft für integrierte Datenverarbeitungssystem mbH, Saarbrücken, 1998.

(Scheer 98a)

Scheer, A.-W.: *ARIS–Vom Geschäftsprozeß zum Anwendungssystem*. 3. Auflage, Springer, Berlin et al., 1998.

(Scheer 98b)

Scheer, A.-W.: *ARIS–Modellierungsmethoden, Metamodelle, Anwendungen*. 3. Auflage, Springer, Berlin et al., 1998.

(Schnotz 94)

Schnotz, W.: *Aufbau von Wissensstrukturen: Untersuchungen zur Kohärenzbildung bei Wissenserwerb mit Texten*. Psychologie Verlags Union, Weinheim, 1994.

(Schütte 97)

Schütte, R.: *Von den Grundsätzen ordnungsmäßiger Modellierung zur Konstruktion von Referenzmodellen*. In: *Forschungsforum '97*, BMBF (Hrsg.), Bonn, 1997.

(Fremdwörterbuch 84)

ohne Autor: *Großes Fremdwörterbuch*. 5. Auflage, VEB Bibliographisches Institut Leipzig, 1984.

(Winter 97)

Winter, A.: *Stellungnahme zum Artikel "Entwicklung eines Referenzdatenmodells für Krankenhäuser"*. In: *Wirtschaftsinformatik* 40 (1998) 4, S. 357 - 358.

Section 3

Electronic Publishing

(Impact on Libraries, Structured Documentation)

Bibliographic database as a support tool for knowledge management

Stefan Kimlicka

Prof. Ing. Stefan Kimlicka, Ph.D.
Department of Library and Information Science, Faculty of Arts,
Comenius University, Gondova ul. 2, 818 01 Bratislava, Slovakia,
e-mail: kimlicka@fphil.uniba.sk

Introduction

Knowledge management is a relatively broad concept. In the milieu of universities and research institutions it includes - among others - the presentation and evaluation of the results of scientific activities. Publication activity and participation at scientific conferences, symposiums and congresses are both very important when it comes to the presentation of scientific results of university teachers and researchers. Via these activities we can follow the directions of research and development on the level of individual scientists, institutions, scientific communities and the state as well. Information about these activities, however, is scattered at many different places around the world. The registration of publication activity is being done at various levels and with a different degree of detail and standardization. It is often required when the evaluation of individual researchers (habilitations, inaugurations, grant applications) or institutions (accreditations) is necessary. Scientific results of researchers are also being recorded in various bibliographic, factual and full-text databases. These databases are, however, created on the principle of a specialization or of a type. So if we want to follow the scientific activities of individuals, groups or nations, we have to search the data at various locations while the completeness of the selection is always a question of luck.

Complexly defined bibliographic database

All the data which is needed for the registration and evaluation of scientific activities from the perspective of published outputs and of the participation at scientific conferences can be stored and processed in a special bibliographic database. A complexly defined bibliographic database contains files and records on 3 levels of distinction:

1. Bibliographic level, for which the records about publications in master file are dominant. Other types of records are derived from them - records about authors (name authorities), records of references, records of source documents (a catalogue of periodicals).
2. Biographic level, where the record of creator (author) in the name authority file is dominant,
3. Scientometric level, where files of periodicals and of scientific actions are important.

The bibliographic database of publication activity must be built upon the appropriate library-information standards, mostly UNIMARC and ISO 2709 (the structure of records for the interchange), AACR2 cataloguing rules and the presentation standards and recommendations of ISBD, STN ISO 690 and Z39.50. Besides that it is necessary to use the authority files and encoding tables to enable the creation of nation-wide registrations and to exchange the records among universities. For having an opportunity of a complex information exploitation of this database it is necessary to create mutual links among individual types of records and files - as it can be seen in fig. 1. These are the links between the records of publications, name authority file, catalogue of periodicals and the authority file - scientific actions. For example the item described in the record no. 11241 has been written by the author registered in the name authority file under no. 061. The article is from a conference - the link to the authorities - scientific actions, record no. 121, and it has been published in the conference proceedings no. 272 (a link to the catalogue of periodicals, record no. 272). Some other articles (Makulova, Kimlicka) are linked to the record no. 121 in the authority file - scientific actions as well. Similarly the links to authority files and references are shown in the record no. 11242. As a result, it is possible to execute various operations and scientometric calculations over this database to produce outputs necessary for the support of scientific activities management. To ensure the highest possible level of completeness in recording the development of a researcher or of a scientific discipline it is also necessary to keep eye on qualification works which quite rarely get listed in the databases but which can contain valuable information. Well-processed and accessible databases of publication activity and of references offer a good picture of the profile and effectiveness of a research institute or university but mainly of individual scientists. Registration and presentation of publication activity and of qualification works on a national level is even more important as a part of the national bibliography system. Its importance here is multiple:

- It presents scientific community on a concentrated national platform within a world-wide system of national bibliographies.
- Processing the „domestic“ researchers' publications published abroad, the database of national bibliography is complemented with the records that are often otherwise unavailable and hence very valuable. It is necessary to mention here the project of national bibliographic registration of qualification works and of grey literature within the internationally recognized and - in the framework of global information environment - implemented system of universal bibliographic control (programmes of UBC and UBCIM).

- Complexly defined bibliographic database may serve as a basis for the biographic research as it can contain the publication activity of a researcher starting with his/her study years.
- It is possible to use this type of a bibliographic database - with the help of appropriate software systems - for various types of bibliometric and scientometric calculations.

Organizing the creation of virtual bibliographic database of science in Slovakia

Registration of scientific, pedagogical and other professional activity via publication output seems to be very unpopular among research workers and university professors. Various kinds of publication activity and references listings are quite regularly or even randomly required by universities and colleges administrations, academy of sciences or research institutes headquarters, ministries, accreditation and grant commissions. Very often these requirements vary in the structure of records and in criteria for evaluation of different types of documents, citations or other kinds of references. These are particularly the categories of document types and their evaluation criteria that are interpreted differently depending on individual evaluation commission. The situation changes according to the orientation of commission members - some of them are only willing to accept the publications in „serious“ scientific journals ranked in Current Contents service meanwhile some of them are able to respect essayistic and popular articles even from the local press. There are some ambiguities in defining the document types and their scientific level as well as in distinguishing types of references. Nevertheless there is no universally valid objective way of evaluating the publication activity and the references. Each scientific branch, each discipline has its own specific features. In spite of all the objections the registration of publication activity and of references represents a meaningful and relatively objective tool of evaluating scientific and pedagogical activity. It is necessary to mention, however, that the data on publication activity itself does not replace the evaluation and, on the other hand, wrong data cannot serve as a basis for a reliable evaluation.

Qualification works are a very special type of documents - habilitation, dissertation, doctoral and diploma theses are the most significant among them and they are usually produced at colleges, universities and institutes of Slovak academy of sciences (SAS). These works are processed and collected very sporadically. Access to them is practically impossible with the exception of nation-widely registered dissertation works that are collected and processed by Univerzitná knižnica (University Library) in Bratislava. These documents are part of the system registering the grey literature. Registration and availability of these works has predominantly a historical importance, but their information and managerial functions are equally important. Historical importance of qualification works' registration lies - besides the documentation of qualification degree - also in the fact that it enables to complete crucial data in biographic researches of well-known personalities. Information function is evident mostly in connection with the role of grey literature. Managerial function consists in

orientation and help with the selection of new dissertation topics and with the evaluation of qualification works.

From the point of view of a complete registration of both types of documents it is important to register and to process all of them, without any respect to their category. After a qualified selection of records from this complex database it is possible to create a virtual bibliographic database of publication activity of scientists and researchers in Slovakia, which can be called SCIENTIA-SK. Potential organizational-functional model of the nation-wide system of processing data on scientific activities is described in fig. 2.

Organizational-functional model has 4 levels and represents the hierarchy of system elements, information flow and locations of providing the outputs and delivering document copies on demand:

- level 1 processing locations of registration and of recording data on qualification works - departments, resp. faculties, where only qualification works are processed, publication activity is processed on a university level
- level 2 processing locations of registration and of recording data on publication activity, scientific actions and qualification works, local databases and locations of document copy distribution on the level of faculty libraries where a central library exists
- level 3 processing locations of registration and of recording data on publication activity, scientific actions and qualification works, local databases and locations of document copy distribution on the level of central libraries
- level 4 nation-wide www servers, resp. locations of central databases.

Local databases of publication activity and qualification works are created on the local level (universities, colleges, faculties, research institutes). These databases can be linked to the databases of electronic copies of documents (full-texts). At this level all categories and types of documents are being registered. The categories of documents express their scientific and professional level, their type specification and territorial impact, resp. importance. Registration databases are created on the nation-wide level as a part of SNB and the records contain a reference to the location databases where the links to electronic copies of documents could be found. On the highest level, the virtual database is compiled by selecting from lower level of nation-wide registration. Hence, virtuality of the bibliographic database SCIENTIA-SK is given by the opportunity of having access from one place to the filtered records from a complex SNB database with consequent links (references) to the records in local databases, their call-numbers, resp. electronic copies in local files, then by linking to the scientific and research actions in local databases and also by having direct links to the references on publications. Registration of all documents of this type without any respect to their present categorization gives us an opportunity to complement or to update the SCIENTIA-SK virtual database in the future.

Virtual view of these databases is also possible from the level of university, resp. college, which has local databases on the faculty level, or from the SAS level when looking at individual institutes. This is also the way to create specialized virtual databases for individual scientific branches or disciplines. The basic precondition for the creation of the system of nation-wide registration of publication activity and of qualification works is the foundation of a cooperative system that would be capable to register and to process bibliographically all these types of documents. For this purpose it is necessary:

- to select an optimal number of fields and subfields in UNIMARC format for the description of publication activity (references to it) and of qualification works,
- to prepare a working methodology for the creation of the above mentioned records in accordance with AACR2 cataloguing rules,
- to prepare a mechanism of interlinking the local databases into central databases on the level of universities and SNB, to ensure the duplicity check-up and the selection of their records to the SCIENTIA-SK virtual database,
- to provide access and maintenance of the SCIENTIA-SK virtual database on the Internet,
- to prepare the competition strategy for the selection of software suitable for the cooperative processing of publication activity (+ references) and qualification works registration,
- to prepare the methodology for the application of STN 01 0195, ISBD and STN-ISO 690 recommendations in the process of editorial preparation of scientific and professional publications in publishing houses,
- to prepare the methodology for the application of STN 01 0195, ISBD and STN-ISO 690 recommendations in the presentation of the records and listings in printed form as well as on the computer and terminal screens within the Internet and Intranet networks.

Creation of a cooperative system of (relatively complete) registration, processing and presentation of publication activity, references, scientific activities and of qualification works at universities and research institutes in Slovak republic represents the principal goal of the **Pro Scientia** project (Kimlicka 1998). The project is based on the application of international standards and recommendations, esp. UNIMARC, ISO 2709, ISBD, AACR2 and Z39.50. It is a part of a larger project of national bibliographic registration of qualification works and publication activity within the Slovak national bibliography in the framework of UBC and UBCIM programmes. Due to the fact that academic libraries involved in the project do not have any personal, financial or coordinative capacities for the solution of problems and tasks of this type, the association of information specialists for the support of science - Pro Scientia - has been founded. Its main activity will consist in concentrating the expertise, obtaining financial resources and creating a pilot project that will be gradually extended to all academic libraries and libraries of research institutes.

The role of libraries in building SCIENTIA-SK virtual database

It is necessary to admit that a frequent need to submit the listings of publication activity and of references, often in various forms and layouts and prepared according to various criteria, is not a pleasant job and usually it takes a lot of time. Registration and processing of qualification works seems to be even more demanding, mostly from the point of view of time and capacity. The solution does not lie, of course, in cancelling the registration of all these documents but in making their processing easier. It is a task for librarians and information specialists working in academic libraries as well as for the preparation team of Slovak national bibliography to find the ways how to do it and how to make the process for the researchers and professors more straightforward.

The complexity of registration and presentation of publication activity and of references can be partially shifted on the librarians' shoulders - after all, they have a qualification for this type of work. It is mostly the question of the basic know-how of a simple document description according to the appropriate standards and the preparation of listings and of other outputs that are also standardized. In order for the librarians to succeed it is important:

- to know the appropriate standards for the document description and the creation of listings and outputs,
- to have a suitable software for the registration of publication activity,
- to divide the competencies and duties among researchers and professors on one side and the librarians on the other, resp. among departments/institutes and the library,
- to create a well-functioning system linked to the library system and the Internet.

The collection of entry data and their proper and reliable input seem to be the most fragile part of the system of publication activity registration (EPC). There are a lot of problems with the complexity of data (to ensure that the documents can be precisely identified) and with the ranking of individual publications, references or scientific activities into categories. If the collection and input of the data is left up to the authors, in most cases the completeness and correctness of data is at stake. This phase should be done by experts - the librarians. On the other hand it is quite natural that the librarians cannot objectively rank the documents into categories of publication activity. This is the domain of authors themselves and the control function can be fulfilled by the scientific community of the discipline represented by e.g. the commission of the scientific board.

Registration and processing of qualification works represents a special problem. A present practice on the academic ground varies enormously. Somewhere the publications are registered and preserved at individual departments, in some cases in the framework of the faculty, in other cases they are just stored without any registration and processing. Reasons of this status quo can be found in the shortage of the storage room for qualification works and in the lack of workforce or of a software suitable for their processing. This

problem can be solved only with financial support from outside of the system and with understanding of its importance from the side of leading figures at the academic institutions.

Software solution for the publication activity and qualification works registration

Based on the experience from the routine run of EPC system at 12 faculties of Comenius University and some other faculties and universities in Slovakia, as well as from the need to approach the international standards it is possible to formulate principal requirements for a good software system for the publication activity and qualification works registration as follows:

- a comfortable input of entry data with the support of authority files (authors, periodicals, scientific and professional actions) and tables (categories of publication activity, categories of references, categories of scientific and professional activities, codes of languages, codes of countries, codes of authors' roles, codes of workplaces, codes of scientific disciplines, codes of document types etc.),
- possibility of conversion of data about employees from the personal module of the faculty/university/institution information system into the name authority file,
- selection of data entry worksheets according to the type of document,
- formal check-up of data during input,
- automatic generation of selected data (date, record number etc.),
- import and export of records in ISO 2709 UNIMARC structure,
- possibility of conversions of databases from other systems of publication activity registration including the conversion of categories (e.g. from the EPC - CDS/ISIS system),
- interlinking of EPC database to the modules of cataloguing, serials management and authorities,
- interlinking of records of publication activity with the appropriate references and with the records of scientific and professional activities (participation at conferences, seminars etc.),
- creation of listings of publication activity, references and participation at scientific, professional and artistic activities sorted according to various aspects (workplace, authors, categories etc.) in accordance with the ISBD rules and the standard ISO 690 as well as according to the requirements of grant agencies, accreditation and evaluation commissions,
- possibility of linking pictures or electronic copies of documents to the records,
- processing of statistical surveys for individual employees as well as the whole institutions,
- application of Z39.50 recommendation to the online catalogue accessible on the WWW-server via Internet with the possibility of searching under a variety of fields and criteria with the help of boolean operators,

- possibility to build a union catalogue of publication activity on the WWW-server (applying Z39.50) on a nation-wide Slovak level and in the network of cooperating bodies (faculties of universities and colleges, scientific institutes of the academy of sciences) and institutions.

System RL-EPCA¹ (Rapid Library - Evidencia publikacnej cinnosti a aktivit) has been tried out in practice and it proved to be a suitable software for the implementation of this type of bibliographic database. This software module can work independently (stand-alone) or as a part of integrated library-information software system Rapid Library. It is based upon the utilization of the UNIMARC structure and of the links that enable to create a relational database and to interlink the records in individual files as it can be seen in fig. 1. Besides a whole range of publication activity and references listings (9 presentation formats and 10 standard outputs) the system provides even the outputs applicable for the support of the scientific activity management. With these functions the software system RL-EPCA ranks beyond the frame of common library-information systems. Provided outputs are as follows:

1. listings of scientific activities of an author or of an institute;
2. listings of articles in periodicals that reach certain threshold of the impact factor for the individuals and for the institutes;
3. an average number of publications per 1 employee (with the possibility of defining the categories of publications for which the calculation applies);
4. impact factor of an author or of an institute for a certain period of time (with the possibility of defining the categories of publications for which the calculation applies);
5. processing of statistical tables for the Accreditation commission (regular evaluations of colleges and universities).

Several tables and encoding tables are built into the software which facilitate and check the input of coded data. Tables of categories were prepared the way so that it would be possible to make the conversions of categories in the whole file even retrospectively, e.g. when the categorization of the Accreditation commission changes. It is also possible - in all of the outputs - to preselect the type of categorization, selection of years' range, threshold of publication categories, authors and institutes, for which the output is produced.

¹ RL EPCA software has been developed and distributed by Cosmotron Systems, Ltd.

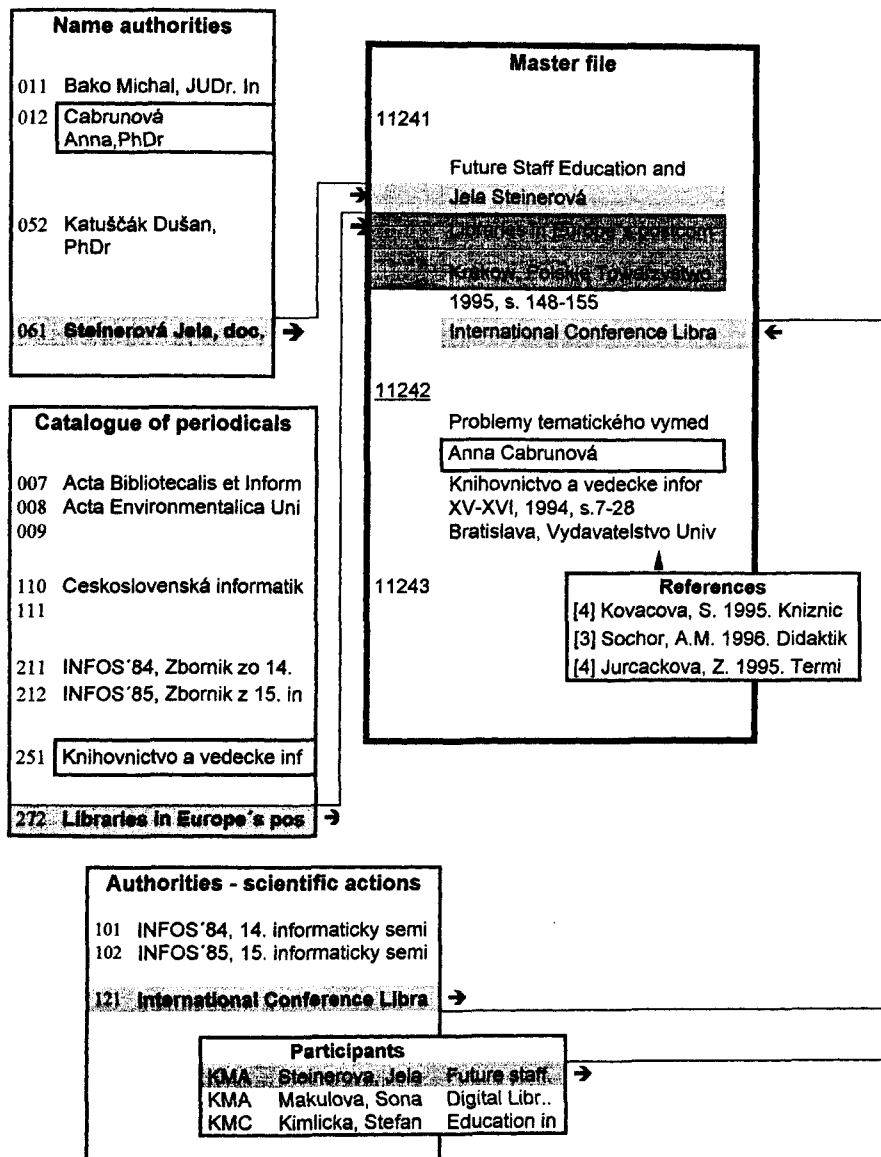


Fig. 1

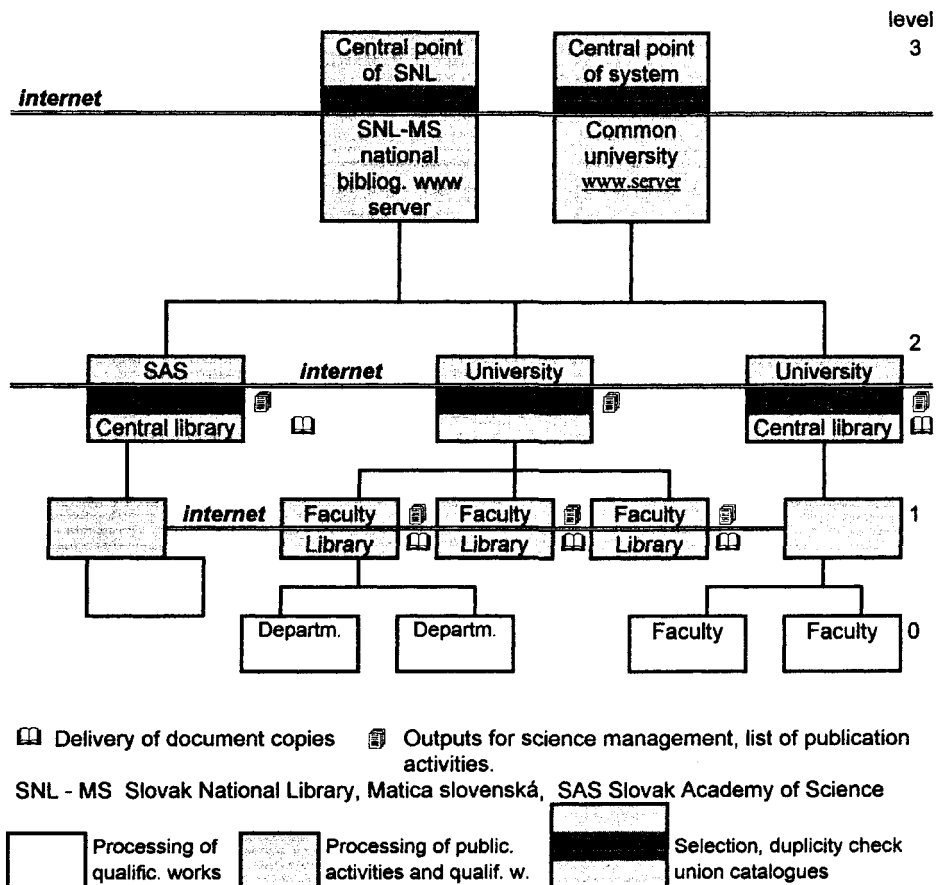


Fig. 2

Summary

Publication activity and participation at scientific conferences, symposiums and congresses are both very important when it comes to the presentation of scientific results of university teachers and researchers. Via these activities we can follow the directions of research and development on the level of individual scientists, institutions, scientific communities and the state as well. Information about these activities, however, is scattered at many different places around the world. The registration of publication activities is being done at various levels and with a different degree of detail and standardization. It is often required when the evaluation of individual researchers (habilitations, inaugurations, grant applications) or institutions (accreditations) is necessary. Scientific results of researchers are also being recorded in different bibliographic, factual and full-text databases. These databases are, however, created on the principle of a specialization or of a type. So if we want to follow the scientific activities of individuals, groups or nations, we have to search the data at various locations while the completeness of the selection is always a question of luck.

All the data which is needed for the registration and evaluation of scientific activities from the perspective of published outputs and participation at scientific conferences can be stored and processed in a special bibliographic database. A complexly defined bibliographic database contains files and records on 3 levels of distinction. 1. Bibliographic level, for which the records about publication in master file are dominant. Other types of records are derived from them - records about authors (name authorities), records of references, records of source documents (a catalogue of serials). 2. Biographic level, where the record of creator (author) in the name authority file is dominant, 3. Scientometric level, where files of serials and scientific actions are important.

To ensure the highest possible level of completeness in recording the development of a researcher or of a scientific discipline, it is also necessary to keep eye on qualification works which quite rarely get listed in the databases but which can contain valuable information. Well-processed and accessible databases of publication activities and references offer a good picture of the profile and effectiveness of a research institute or university, but mainly of individual scientists. Registration and presentation of publication activities and qualification works on a national level is even more important as a part of the national bibliography system. Its importance here is multiple:

- It presents scientific community on a concentrated national platform within a world-wide system of national bibliographies.
- Processing the „domestic“ researchers' publications published abroad, the database of national bibliography is complemented with the records that are often otherwise unavailable and hence very valuable. It is necessary to mention here the project of national bibliographic registration of qualification works and grey literature within the internationally recognized programmes of UBC and UBCIM.

- Complexly defined bibliographic database may serve as a base for the biographic research as it can contain the publication activity of a researcher starting with his/her study years.
- With the help of appropriate software systems it is possible to use this type of a bibliographic database for different types of bibliometric and scientometric calculations.

The structure and links within a complex bibliographic database and the Rapid Library - EPCA (Evidencia publikačnej činnosti a aktivít - Registration of publication activities) software are discussed in this study. The main function of this software is the processing of the data for the purposes of research activities management at universities and research institutions. Beside standard types of library-information outputs it provides the possibility to create various bibliometric and scientometric outputs. This software is being implemented within Pro Scientia project, main task of which is to build up a cooperative registration system of scientific publication activities on a national level in Slovakia.

Bibliographic references

KIMLICKA, Stefan. 1997. Informacny system o publikacnej cinnosti - sucast integrovaného univerzitného informacného systému. In: *UNINFOS'97 : Univerzitne informacne systémy*. Kosice : Technická univerzita, 1997, p.28-36.

KIMLICKA, Stefan. 1998. Evidencia publikacnej cinnosti a kvalifikačných prác - základ virtualnej bibliografickej bazy dát SCIENTIA-SK. In: *Bulletin CVTI SR*. Vol. 2, 1 (1998), p.3-11.

KIMLICKA, Stefan. 1998a. Programový systém RL-EPCA na podporu riadenia vedy na univerzite. In: *UNINFOS'98 : Univerzitne informacne systémy*. Nitra : Slovenská poľnohospodárska univerzita, 1998, p.44-47.

Information Research: a case study in the free electronic publication of research

T.D. Wilson

T.D. Wilson, Ph.D.
Research Professor in Information Management
Department of Information Studies
University of Sheffield
t.d.wilson@shef.ac.uk

Contents

1. Introduction - the rise of the electronic journal
2. The history of *Information Research*
3. The aims and functions of *Information Research*
4. Production and usage
 - 4.1 Economics
 - 4.2 Usage
5. The future of *Information Research*
6. Conclusion
7. References

Summary

Describes the development and current status of the electronic journal, *Information Research*, which began as a print-based Departmental newsletter and which now publishes original research papers. The discussion covers the rise of the e-journal concept and its implications for academic institutions, as well as the future of e-journals and the possible development of *Information Research* as a collaborative educational and research venture.

1. Introduction - the rise of the electronic journal

The concept of the digital library has now become almost commonplace - conferences act as fora for the latest research, the National Science Foundation

in the USA and the Higher Education Funding Councils in the UK have research programmes in various aspects the digital library, journals advertise special issues, and new journals arise. Indeed, it seems almost as though the sole topic of professional writing has become the digital library.

Central to the concept of the digital library is the fact of the digital *resource*, since digital libraries cannot exist without being able to provide access to digital resources, and one of the most significant developments over the past four or five years has been the rise and continued rise of the electronic journal as a serious medium of scientific communication.

That rise is indicated by the size of the archive at *NewJour*, a Web-site which lists new electronic journals and from which an e-mail service of announcements is run. *NewJour* began life in 1993 as part of the Association of Research Libraries directory of electronic resources. In early 1995 there were 250 items in the archive, there are now (24th August 1998) 6,255 items. Not all of these are serious, academic journals: *NewJour* lists newsletters, popular magazines, trade journals and, indeed, any item that can be described as a periodical publication. However, many are serious journals and a particular feature, as many will be aware, is the rapid increase in the number of priced journals, published by the major journal publishers.

The fact that the publishers are rapidly making their journals available in electronic form is not, for me, the interesting phenomenon — clearly, they must do so if they are to survive this electronic revolution. More interesting is the fact that many of the titles in *NewJour's* archive are *free* electronic journals in many different disciplinary fields. Some are archival projects, like the *Jahrbuch über der Fortschritte der Mathematik* (<http://www.emis.ams.org/projects/JFM/>), which is a searchable database of the contents of the *Jahrbuch* from 1868 to 1943, others are in relatively new disciplines or professional areas, where the volume of demand for print journals may be insufficient to support a commercial publication. For example: *MC Journal: The Journal of Academic Media Librarianship* (<http://wings.buffalo.edu/publications/mcjrnl/>); *Journal of Australasian Graphics Imagery* (<http://www.cs.curtin.edu.au/jagi/>); *Journal of Buddhist Ethics* (<http://jbe.la.psu.edu/>); *Journal of Digital Information* (<http://jodi.ecs.soton.ac.uk/contents.html>); and the *Journal of Electronic Publishing* (<http://www.press.umich.edu:80/jep/>). Others are, perhaps, on the fringe of 'normal' academic research, such as the *Journal of Applied Missiology* (<http://www.acu.edu/academics/missions/jam.htm>), which is devoted to the study of Christian missionary work.

Of course all of these examples are from only the „J“ section of the *NewJour* archive, a survey of the entire file would reveal many more.

Why this massive increase in the number of electronic journals, and why are so many published free of charge? The answer to this has to do with the nature of academic research, with the costs of publishing, and with the complex relationship between the academic, his or her parent institution, the research

community, and the patterns of scholarly communication. First, in most countries it is anticipated that the university teacher will also conduct and publish research – it is part of the contract. Consequently, the demand on print-on-paper publishing (POPP) sources in many disciplines is high and, even if accepted through the refereeing process, a paper is likely to take months to reach the academic's peers. POPP is slow. Electronic publication, on the other hand can be rapid and the only real bottle-neck in the publishing process is the refereeing stage and even this is generally faster than in POPP, since the document can be transferred electronically and commented upon by e-mail. If the editor controls quality, or if the board of referees is based within a single institution, the refereeing process can be very rapid.

The costs of POP publishing are also high – certainly to the consumer. It is a well-established fact that the costs of academic journals have increased considerably ahead of the increases in the costs of living. Beginning a print journal requires the academic to convince a publisher (or his/her own institution) that the journal can break-even in a relatively short period of time, say, three to four years and, for a commercial publisher, that it can make a profit not long thereafter. If your field is a very small one and very esoteric, it is rather unlikely that a print publication is a realistic opportunity for a publisher. On the other hand, the start-up costs for an e-journal are either small or non-existent since, in the academic world, the intending editor/publisher will almost certainly have access to an institutional or departmental Web-server and the costs of putting files on the server are minute in comparison to the costs of producing printed pages. The e-journal editor is also saved virtually all production and, importantly, distribution costs for the journal. Production costs are low, amounting to the time needed by the academic to edit HTML documents (and little editing may be needed if style sheets are produced for contributors to work to) and set up the hyperlinks for the next 'issue' of the journal. Naturally, the load depends upon the frequency of publication and a quarterly issue will be less time-consuming than a monthly.

Turning to the patterns of scholarly communication, the original pattern of such communication was very much through personal communication and attending the meetings of the early scientific societies, such as the Royal Society. The *Proceedings* of such meetings were mainly to advise members of what had gone on at the last meeting. As expansion of the scientific enterprise took place, the numbers of members increased, more and more could not attend all meetings, and the *Proceedings* became a vehicle for not only reporting on events, but also for publishing papers that had *not* been presented at meetings. As membership grew, the number of non-members also interested in the *Proceedings* increased and distribution became more of a problem for the scientific society – at this point the commercial publishers enter the scene, because they already had distribution chains for their other publications, as well as the infrastructure to develop. We can argue that the emergence of the free e-journal is a return to the ideals of scientific communication – the free distribution of the results of research among those seeking to advance the discipline.

2. The history of Information Research

Information Research: an electronic journal is now entering its fourth year of publication. It had its origins in a small newsletter, *CRUS News*, which reported on the work of the Centre for Research in User Studies. *CRUS News* came to a natural end on the cessation of the Centre's contract and, although there was considerable interest in the work of the Centre and associated research (we continue to receive enquiries some 10 years after it ceased to function), it was not possible to maintain *CRUS News* without the researchers who had carried out the related work. However, it was decided to expand the role of *CRUS News* and to use it (under the title *Information Research News* or *IRN*) to publish working papers on research in the Department.

IRN was published from 1990 to 1997 in paper form and, from April 1995 (under the title *Information Research*), in electronic form. During 1994/95 it became evident that continuing to publish *IRN* on paper did not make economic sense: when all costs were included, the income from the small number of subscriptions barely covered production costs, and in the course of 1995 it became evident that the electronic version, *Information Research*, was reaching many more readers than the paper version had ever done and, therefore, in 1997 the decision was taken to publish only the electronic version.

3. The aims and functions of Information Research

Information Research was originally designed to publish working papers, rather than fully elaborated papers that could be exposed to referees. However, from the response to the papers, in terms of the usage data as well as e-mail messages, it became evident that the journal was accepted as a 'normal' academic journal. Over the first three years, therefore, the aims of the journal have changed from publishing mainly working papers based on research carried out in the Department of Information Studies, to publishing working papers, invited papers (known as 'guest papers') from outside the Department, and, most recently, fully refereed papers, including, in the current issue, papers presented at a recent conference (this was done at the request of and with the full collaboration of the conference organizers). We aim to publish research papers of a quality that would be acceptable in a print journal and will increasingly rely on a Board of Referees to ensure that this is the case.

Information Research also has two other functions – one internal to the Department, the other external. The internal function is that it serves as a valuable public relations tool for the Department and, for example, attracts applications from potential Master's and Ph.D. students. With a world-wide audience, the PR value is considerable. The external function is that it provides users with free access to research papers and it is evident that many of those who register to receive update information, especially from Third World countries, find this particularly useful, since economic factors limit their ability to use print publications.

4. Production and use

4.1 Economics

In its original form, the economics of *Information Research* were relatively trivial: the Department used the University's Web server and, therefore, no capital costs were involved, one person (the author) converted original papers into HTML format and passed them to the Department's Computer Manager for uploading to the Web-site. In all, the work-load was probably not more than two person-days a month at a cost of, say, £150.00. Since the Web-site was run and maintained by the University's central computer services, there were no readily identifiable maintenance costs separately attributable to *Information Research*. Material costs were similarly insignificant, since the papers were already in electronic form and all editing and conversion was done on the electronic versions.

Editorial work of one kind or another is generally accepted as falling within the normal work of an academic member of staff in a British university, and in many others around the world. We are encouraged to serve on Editorial Boards, to act as referees, and to serve as Editors. These tasks are seen as part of the involvement in research. I serve on several Editorial Boards and have been the Editor of two journals in the field, both of which I started from the stage of preliminary negotiation with the publisher through to production and development. In a real sense, Sheffield University subsidises the activities of publishers not only through my involvement but, of course, through the involvement of many more. It would seem curious if it was to seek to prevent similar involvement in electronic journal production from within the institution itself! If my, and others', involvement in editorial work is provided by the university without cost to the publisher, it is hardly likely that it would like to see a return on my work as editor and producer of an electronic journal, other than the kudos that might accrue to the institutions from the very existence of the journal. Therefore, both for commercial publishers and for the institution itself, this kind of involvement is seen as cost-free by the university.

The production of a *free* electronic journal has certain benefits in terms of other costs: it is not necessary to keep records of subscribers, no accounts need to be kept, no letters have to be sent to subscribers urging them to renew, we pay no commission to other businesses to act as agents, we have no offices within which all of these activities are carried out and employ no staff to perform them. These overhead costs are quite significant for both paper and electronic journals that are charged: for example, Fisher in a paper to the Conference on Scholarly Communication and Technology in 1997, provided the following information, comparing the overhead costs associated with and 'issue' of the *Chicago Journal of Theoretical Computer Science* (CJTCS) with those of an issue of the paper journal *Neural Computation* shows:

	CJTCS	NC 8:5
	\$	\$
Journals Department		
Production	8,000	1,00
Fulfillment Cost Per Subscriber	108	1
General and Administrative	31,050	2,30
Digital Projects Lab		
Staff	200	--
Hardware and Software	5,000	--
Total Overhead Per Subscriber	44,358	3,30
OH costs per page published	182	14

The author notes that:

From the comparison between CJTCS and Neural Computation, it seems that the direct costs of publishing an electronic journal are substantially below that of a print journal with comparable pages. The overhead costs, however, are much higher – 1240% higher in this case... (Fisher, 1997)

She goes on to note that this disparity in overhead costs is largely a function of the amount of content published by CJTCS in the 18 months over which overhead was calculated, compared with the distribution over 12 issues of the print journals, NC. In other words, like was not being compared with like. If you do not have subscribers, but only readers, a significant overhead costs is clearly removed, at least the \$31,050 of administrative charges.

On the other hand, if *Information Research* has cost, let us say, £600 since April 1st this year (and it is probably much less), in which time we have published two issues and have had (at 21st August) a total of 4,520 hits, then the cost per hit is only 13 pence (or about \$0.21) - to recoup costs from the registered users (283), and assuming that my cost calculations are 100% out, we would only have to charge about £14.00 a year - however, we could not do this, since the University will not let us charge less than £25.00 for anything because it is not worth charging less in terms of the associated accountancy costs! These numbers do not appear to be too far out, since, for example, the individual subscription cost of CJTCS is only \$30.00 (£18.00) a year.

4.2 Usage

Information Research originally had a counter attached to it, which recorded all „hits“. However, this was abandoned as a result of a change in policy for the University server on which the journal is mounted - at the time it was abandoned, the journal was averaging about 800 hits a month. Since Volume 3

number 4, a free counter service (available from NedStat at <http://usa.viewstat.nedstat.net/>) has been used and counters have been provided for individual papers through the Link Exchange Fast Counter service at <http://fastcounter.linkexchange.com/>.

The statistics of usage since 1st April 1998, show (at 3rd September 1998) that the top page of the journal ([ircon.html](#)) has had 4875 hits - or approximately 1,000 a month. This, of course, is not the true total - since users will be going directly to specific papers, as a result of search-engine searches and back to the same papers to re-read or pick up references, or whatever. Users come from eighty-five Internet domains, plus 20% from unknown domains. The highest-using domains are shown in Table 1

The fact that these thirteen domains cover 62% of usage and that 20% is from „Unknown“ means that the remaining seventy-two domains have very small usage, accounting for only 18%, or 0.25% each, which, at the 21st August 1988 means about 10 hits each since 1st April. The full list of domains is given in the Appendix Table A.

Those hits that could be assigned to a geographical region were allocated as shown in Table 2. There is clearly a relationship between the regions that

1.	United Kingdom	17.72 %
2.	US Commercial	10.84 %
3.	US Educational	8.67 %
4.	Network	7.39 %
5.	Australia	4.05 %
6.	Germany	2.50 %
7.	Canada	2.21 %
8.	Spain	1.95 %
9.	Hong Kong	1.73 %
10.	Malaysia	1.57 %
11.	Singapore	1.28 %
12.	Finland	1.15 %
13.	Korea (South)	1.08 %
Total		62.14%

Table 1: Internet domains of
Information Research users

Europe	31.17%
North-America	23.63%
Asia	10.01%
Australia	4.64%
Africa	0.93%
South America	0.91%
Central America	0.69%
None (.net, .org)	8.00 %
Unknown	20.03%

Table 2: Geographical distribution of
hits by region

dominate and the spread of effective Internet connections in those regions and the dominance of Europe and North America is, therefore, to be expected.

The busiest time of day is between 16:00 and 18:00 GMT - mainly, it seems, because of the volume of US traffic at that time - readers in Europe who want easy access, therefore, would be advised to use the site in the morning. The same pattern exists, of course, for on-line services and it is well understood that if you want easy access to, say, Dialog databases, it is best to do your searching before the USA wakes up. Similarly we can identify Wednesday as the busiest day of the week, with Friday being the least busy - however the

difference between them is not great: 18.9% on a Wednesday and 14.1% on a Friday. Of course, there is always a surge in demand whenever a new „issue“ is announced.

Hits are a relatively crude way of assessing usage, but one of the few ways we have when producing electronic journals. We also suggest, however, that readers should register to receive information on the timing and contents of new issues and, at 21st August, there were 283 registered readers, 199 of whom were distributed as in Table 3 (the full distribution is shown in the Appendix - Table B). That is, 70% of the registered readers came from 12 countries out of the total of 53 - once suspects that the Bradford distribution, so well known in bibliometrics, exists here also.

UK	67
USA	50
Australia	15
Germany	9
India	9
Canada	8
Singapore	8
China	7
Ireland	7
Malaysia	7
Brazil	6
Spain	6

Table 3: Geographical distribution of registered readers

5. The future of Information Research

I have already described how, since 1995, *Information Research* has moved towards a fully peer-reviewed research journal and I anticipate that the movement will continue and that, over the next three years or so, the journal will be publishing mainly refereed papers. We are also exploring the possibility of making *Information Research* a collaboratively-produced journal, and we are having discussions to that end with three other departments in different parts of the world. These other departments will appoint a member of staff to act as regional Editor for the journal, seeking contributions, and preparing the papers for submission to the Sheffield server. In time, if demand on the server grows, the site may be mirrored in these institutions – this will give some problems in terms of monitoring demand, but I am sure that the problems can be overcome.

The collaborative model is a very interesting one from the point of view of the scholarly journal because it distributes whatever workload there is over several institutions and allows for more institutions to join the consortium if the workload grows. The working paper contributions, which usually involve collaboration between students and teachers will also serve to stimulate such collaboration

within the partner institutions and the regional Editor might also build the HTML conversion process into teaching to give students some feel for the genuine production issues of electronic journals. We might, for example, hand over the entire production of an 'issue' to one of the partners to use as a project for a class on electronic publication.

6. Conclusion - the future of the free scholarly journal

The growth of free electronic journals depends upon a number of factors:

- a) the extent to which researchers in new and/or multidisciplinary fields find it difficult to publish in core journals and difficult to persuade commercial publishers of the existence of a big enough market for a print journal;
- b) whether speed of publication and world-wide exposure will outweigh the perceived value of the citation of papers in print journals;
- c) how quickly electronic journals come to be covered by the citation indexes;
- d) how much longer academic institutions will be prepared to tolerate the present *uneconomic* situation in which they are, effectively, subsidising the profits of commercial publishers;
- e) whether the scientific societies will find new sources of income through gaining sponsorship of free electronic journals to replace the income they get by contracting commercial publishers to produce their journals.

All of these are big questions and papers could be written on each of them, but I believe that the logic and economics of the free publication of scholarly research will be overwhelming and that we shall see a return to the ethos of the free interchange of knowledge in a genuine community of scholars. We no longer need to depend upon the technology that has served research well for the past 350 years, and the new technology offers not only speedy publication but multimedia publication, which is very attractive for many fields. The new technology will give rise to new models of the process of scientific communication and, for academic institutions, new models of the research dissemination activities of their staff members. When all the circumstances are right, the trickle of new, free e-journals will become a flood — and will cause new problems for librarians and users. One thing we can be sure of: just as librarians have coped with many of the problems of POP publication, so they will cope with the different problems of electronic publication.

7. References

[Fisher, J.H. 1997]

Fisher, J.H. *Comparing Electronic Journals to Print Journals: Are There Savings?* Paper delivered to the Conference on Scholarly Communication and Technology, Emory University, April 24-25 1997. Available at: <http://www.arl.org/scomm/scat/fisher.html> Site visited 25th August 1998.

[Information Research 1998]

Information Research: an electronic journal. Available at: <http://www.shaf.ac.uk/~is/publications/infres/ircont.html> Site visited 25th August 1998

[NewJour 1998]

NewJour: electronic journals and newsletters. [Archive site] Available at <http://gort.ucsd.edu/newjour/> Site visited 25th August 1998.

Informations- und Wissensmanagement durch strukturierte Dokumente

Jiri Panyr, Werner Zucker

Siemens AG, Abt. ZT IK 1
Otto-Hahn-Ring 6, 81739 München
e-mail: jiri.panyr@mchp.siemens.de / werner.zucker@mchp.siemens.de

Inhalt

1. Einleitung
2. Probleme und Ursachen bei der bisherigen Vorgehensweise
 - 2.1 Seiten- und Dokument-Metapher
 - 2.2 Dokument-Management-Systeme
 - 2.3 Information Retrieval / Informationsfilterung
 - 2.4 Executive Information System
 - 2.5 World Wide Web / Elektronisches Publizieren
 - 2.6 Konsistenzproblematik / Dokumentenaustausch (Protokolle)
3. Randbedingungen
 - 3.1 Kundenforderungen
 - 3.2 Qualitätsmanagement durch ISO 9000
4. Optimierungsbedarf bei der Dokumentation
 - 4.1 Inhaltsorientiertes Strukturieren von Dokumenten und Information
 - 4.2 Konsistenzproblematik
 - 4.3 Auffinden von Information
 - 4.4 Durchgängigkeit im Prozeß durch Dokumentenaustausch
 - 4.5 Automatisiertes Erzeugung von Dokumentation
 - 4.6 Single Source Publishing /Single Source Information
5. Weltweite Aktivitäten in Auswahl
6. Resume

Zusammenfassung

Technische Dokumentation ist ein wichtiger Bestandteil komplexer technischer Systeme. Sie ist als Informationsressource sowohl für den Kunden als auch für den Hersteller von entscheidender Bedeutung. Wesentlich ist dabei ihre inhaltliche Konsistenz, Austauschbarkeit und Wiederverwendung. Je umfangreicher

und komplexer die technische Dokumentation ist, um so wichtiger ist ihre inhaltliche Strukturierung. Während sich die Strukturierungstechniken (und die Trennung von Inhalt und Layout) im Dokumentationsbereich zunehmend durchsetzen, fehlen bei der Anwendung strukturierter Dokumente meist Werkzeuge und Verfahren, die diese Strukturen auch nutzen. In der Arbeit wird auf diese Problematik hingewiesen. Zunächst werden die Probleme mit dem bisherigen präsentations- und seitenorientierten Paradigma aufgelistet, dann auf die Verbesserungsmöglichkeit durch die Nutzung der Strukturierung hingewiesen. Schließlich werden einige weltweite Aktivitäten im diesem Bereich genannt.

Abstract:

Technical documentation is an important part of complex technical systems. As an information resource, it is of critical importance for the customer as well as for the producer. Consistence in content, interchangeability and reuse is essential. The more extensive and complex the technical documentation the more important is the content structuring. While techniques of structuring (and the separation of contents and layout) in the documentation area are more and more accepted, tools and methods in application of structured documents are missing that utilize these structures. This paper focuses on these problems. First the problems with the current presentation and page oriented paradigm are listed, then the possibility of improvement through using the structuring is explained. Finally a survey of world wide activities in this area is presented.

1. EINLEITUNG

In der Informationsgesellschaft von heute kommt der Bereitstellung, Verarbeitung und dem Austausch von Informationen eine überragende Bedeutung zu. Dennoch wird der Effizienz und der Effektivität beim Umgang mit ihr nur geringe Bedeutung beigemessen, besonders dann, wenn Information in Dokumentform vorliegt. Dann wird zum Teil erheblicher Aufwand für die Darstellung von Dokumenten (d.h. für Layout und Typographie), nicht aber für ihre inhaltlich logische Strukturierung geleistet. Das schlägt sich nieder in den zum Einsatz kommenden DTP-Werkzeugen und -Formaten und findet im Boom der Web-basierten Informationsdistribution seine erneute Bestätigung.

Auch dem Aspekt der wiederverwendbaren und weiterverarbeitbaren Information wird, von Ausnahmen abgesehen, nach wie vor nur eine geringe Aufmerksamkeit gewidmet. Besonders die Dokumentation von technischen Systemen wird im Wertschöpfungsprozeß eher als notwendiges Übel betrachtet. Anstelle von informationsorientierter Verarbeitung tritt daher die präsentationsorientierte Verarbeitung. Anstelle semantischer Informationseinheiten treten visuelle Einheiten. Die Strukturierung erfolgt mittels Layout auf der grafischen Ebene, nicht aber auf der logisch semantischen Ebene.

Technische Dokumentation ist ein wichtiger Bestandteil komplexer technischer Systeme. Sie ist als Informationsressource sowohl für den Kunden als auch für den Hersteller von entscheidender Bedeutung. Wesentlich ist dabei ihre inhaltliche Konsistenz, Austauschbarkeit und Wiederverwendung. Je umfangreicher und komplexer die technische Dokumentation ist, um so wichtiger ist ihre inhaltliche Strukturierung. Während sich die Erkenntnis der Strukturierungsnotwendigkeit und die Trennung von Inhalt und Layout zumindest im Dokumentationsbereich zunehmend durchsetzt, fehlen bei der Handhabung strukturierter Dokumente meist Werkzeuge und Verfahren, die diese Strukturen auch nutzen. Mit anderen Worten: Die inhaltlichen Strukturen von Dokumenten sind zwar vorhanden, viele Anwendungssysteme (z.B. Information Retrieval, Suchverfahren in WWW, Desktop Publishing Systeme, Executive Information Systeme) setzen in der Regel weiter auf das alte gesamtdokumentausgerichtete und präsentations- bzw. seitenorientierte Paradigma auf.

2. Probleme und ihre Ursachen bei der bisherigen Vorgehensweise

Zum o.g. älteren Paradigma werden im folgenden einige ausgewählte Beispiele aufgelistet.

2.1 Seiten- und Dokument-Metapher

Die bisherigen Textverarbeitungssysteme (wie z.B. Word etc.) und auch Desktop Publishing Systeme (DTP-Systeme, wie z.B. Interleaf) arbeiten primär präsentationsorientiert, d.h. die dokumentarische Bezugseinheiten sind ganze Dokumente, während die Präsentationselemente die Dokumentseiten sind. Es wird in diesem Kontext auch über Dokumenten- bzw. Seiten-Metapher gesprochen. Im Vordergrund steht bei allen Ansätzen nicht der Inhalt, sondern primär das Layout. Dies führt u.a. zu folgenden Konsequenzen:

- **Seitenorientierte Navigation / Logische Einheiten nicht als Sprungziel**
Bei der Navigation in Dokumenten wird unabhängig davon, ob Verweise präzise auf eine bestimmte Textstelle zeigen, von einer Reihe von Viewern (wie die z.B. von Adobe, Frame, Interleaf) nach dem "Anklicken" lediglich der Anfang der Seite, die das Querverweisziel enthält, angezeigt. Eine präzisere Positionierung innerhalb einer Seite ist in einem der o.g. Viewer nicht möglich. Heutige Beschreibungsformalismen für Dokumente erlauben es nicht, logische Einheiten als Sprungziel zu verwenden. Dies führt einerseits zu den o.g. Problemen bei der Navigation, andererseits führt dies zu semantischen Ambiguitäten, wenn beispielsweise die unmittelbare Umgebung eines Sprungzieles bei der Bearbeitung gelöscht wird. Es ist dann nicht klar, ob der Link noch gültig ist oder nicht.

- **Unzureichende Methodik und Systematik**

Obwohl auch DTP-Systeme gewisse Strukturierungshilfe (wie z.B. Komponenten) bieten, sind diese so gut wie nicht genutzt, da eine logische Systematik (Vorschrift) und Methodik bzgl. ihrer Nutzung fehlt. Komponentennamen werden willkürlich vergeben. Komponenten haben auch in gleich garteten Dokumenten oftmals eine unterschiedliche Benennung und Gestaltung (Eigenschaften), obwohl sie die gleiche Position und Funktion im Text innehaben.

- **Informationsverlust bei Konvertierung**

Die Konvertierung der von DTP-Systemen verwendeten Formate (wie ps, pdf, ...) führt zu einem erheblichen Informations- und Layoutverlust. Die Erhaltungsfunktionalität basiert in der Regel noch immer auf Formaten (wie z.B. DIF, DCA, RTF, MIF), die inhaltliche Erweiterungen bei der Konvertierung nicht berücksichtigen.

Die o.g. Darstellungsorientierung unter gleichzeitiger Vernachlässigung der inhaltlichen Aspekte beim Textaufbau führt zu Reihe von Problemen bei der inhaltlichen Erschließung von Dokumenten und Wiedergewinnung von Informationen, sowie auch bei ihrer Präsentation in einer Hypertext-Umgebung. Die bis vor kurzem in diesem Bereich dominierende textorientierte Künstliche Intelligenz (KI) hat diese konservative Metapher mehr oder minder unterstützt und keine neue Paradigma befördert.

2.2 Dokument-Management-Systeme

Konventionelle Dokument-Management-Systeme verwalten Versionen und Varianten, kontrollieren die Zugangsberechtigung und die Freigabe, verfügen über Änderungs- und Revisionsmanagement und meist geringfügige Workflow-Funktionalität. Kennzeichnend für alle Systeme ist das speicher-orientierte Management von Dokumenten. Es werden Files verwaltet jedoch nicht Information. Die kleinste managbare Einheit ist das File und nicht die Information. Der Zugriff auf Information in Dokumenten erfolgt nicht über die logischen Struktur der Dokumente sondern über die Attribute von Files. Information Retrieval findet dokumentenbezogen statt und kann nicht auf bestimmte Teile fokussiert werden. Jegliche Funktionalität von Dokument-Management-Systemen ist File-orientiert. Dokument-Management-Systemen sind somit eher komfortable File-Management Systeme und weniger Informations-Datenbasen.

Desweiteren nutzen konventionelle Dokument-Management-Systeme die Erkenntnisse aus der DB-Technologie so gut wie nicht (Beispiel: Datenstrukturierung, Redundanzfreiheit, zentraler Zugriff, benutzerspezifische Views, Normalisierung), d.h. sie sind lediglich unvollkommene Datenhaltungsinstrumente. Es bestehen Schwierigkeiten z.B. bei Wiederverwendung und Ähnlichkeitsnutzung dieser Dokumente. So wird z.B. auch Ähnlichkeitsplanung bei CAD/CAP-Dokumenten nicht unterstützt, da die im DMS abgelegte Dokumente

de facto nicht präzise recherchierbar sind. Die im DMS abgelegten Dokumenten können inhaltlich nur rudimentär beschreiben, d.h. erschlossen. Inhaltliche Extraktion ist bisher nur durch unvollkommene syntaktisch-semantische Werkzeuge möglich. Obwohl immer eine Aufgabe der text-orientierten KI-Forschung, sind die Lösungen bisher unbefriedigend.

Bei allen Anwendungen (z.B. Entwicklungsdokumentation) ist eine elektronische Übernahme von Dokumententeilen wünschenswert, findet jedoch hauptsächlich wegen technischer und organisatorischer Inkompatibilität nicht statt.

2.3 Information Retrieval / Informationsfilterung

Die präsentationsorientierte Dokumentgestaltung verursacht bei der inhaltlichen Erschließung und bei der gezielten Wiedergewinnung (Retrieval) von Informationen eine Reihe von Problemen, insbesondere bei längeren, inhaltlich nicht weiter strukturierten Dokumenten. Nicht ohne Grund sind bis jetzt die besten Retrievalergebnisse mit Dokumenten in Form eines Abstracts erzielt. Eine automatische oder zumindest semi-automatische Erstellung eines signifikanten Textextrakts oder eines redundanzfreien Textabstrakts wurde bisher vor allem durch ungenaue auf den ganzen Text anwendbare syntaktisch-semantische Werkzeuge versucht. Im weiteren werden einige Probleme kurz skizziert:

- Große Probleme gibt es mit der automatischen Indexierung bei Dokumenten, die als nicht weiter differenzierbare Textmasse präsentiert sind. Falls Strukturierung im Dokument vorhanden ist, dann lediglich in Form von Layout. Bei größeren Dokumenten ist daher eine sinnvolle Selektion der bedeutungstragenden Wörter kaum möglich. Semantische Textanalysen allein sind nicht ausreichend. Daher ist vonnöten, die zu erschließende Texte schon bei ihrer Erstellung einer inhaltlichen Gliederung zu unterziehen, mit konsequenten Trennung von Inhalt und Layout.
- Die unvollkommene inhaltliche Erschließung von Dokumenten verursacht auch große Probleme bei ihrer Wiedergewinnung (Retrieval, Recherche), wie z.B.:
 - Suche ist nur im ganzen Dokumenten möglich bzw. in ihren strukturierten Teilen (Titel, Autor, Verlag etc.).
 - Bei größeren Texten entstehen viele überflüssige Treffer, weil keine weitere inhaltliche Textdifferenzierung (s.o. Indexierung) möglich ist.
 - Relevanzfeedback-Verfahren (interaktive Retrievalstrategien) operieren lediglich auf den Beschreibungen von ganzen Dokumenten (bzw. ihrer Abstracts). Sie ermöglichen zwar eine Recall-Erhöhung, liefern jedoch gleichzeitig auch viele Ballastdokumente.
- Ähnliche Probleme entstehen auch bei der Informationsfilterung, die ebenfalls nicht komponenten-orientiert ist, d.h. es werden zu viele Ballastdokumente ausgewählt. Die nutzergerechte Verteilung der gefilterten Informationseinheiten (SDI) wird dadurch erschwert.

Die Entwicklung von neuen Methoden im Bereich der Information-Retrieval-Systeme hat infolge der nichtrealistischen Versprechungen der KI über Jahre stagniert. Die Konzepte derzeit eingesetzter Systeme (auch der neueren) basieren in der Regel immer noch auf den Ansätzen vom Ende der 60iger bis Anfang der 70iger Jahre. Hier helfen nicht mehr einige Änderungen im Detail, sondern eine Neukonzeption der Dokumenterstellung und -erschließung. Projekte, die das mit einem realistischen Zeitrahmen und vernünftigen Ressourcen ermöglichen, sind bisher nicht in Sicht.

2.4 Executive Information System

Seit Ende der 80er Jahre setzt sich zunehmend eine neue Variante der Management-Unterstützungssysteme durch, die sich unter dem Begriff *Executive Information Systems* (EIS) zusammenfassen läßt. Diese DV-Systeme - in BR Deutschland auch als *Führungs-Informationssysteme* (FIS) oder *Chef-Informationssysteme* (CIS) bezeichnet - sind in der Nähe von Berichtssystemen anzusiedeln, wobei sie jedoch aufgrund ihrer großen Flexibilität und Funktionalität (individuelle Drill-Down-Technik und Ausnahme-Berichtsgenerierung, Trendanalysen, Einbindung von e-mail etc.) über die Grenzen traditioneller (DV-gestützten) Berichtssysteme hinausgehen und vom Top-Management selbst benutzt werden sollen. In Abgrenzung zu älteren Entscheidungsunterstützenden Systemen werden EIS nicht für die Lösung schlecht strukturierter Entscheidungsproblematiken herangezogen, sondern sind vielmehr personenbezogen auf bestimmte Entscheidungsträger und deren Informationsbedarf ausgerichtet. Sie umfassen jene aufbau- und ablaufsorganisatorischen Regelungen, Führungsgrößen und -methoden, die zur Initiierung, Planung, Analyse und Kontrolle von Geschäftsprozessen im Unternehmen benötigt werden. Damit EIS bedarfsgerechte Unterstützung (des Top-Management) leisten können, muß ein hoher Integrationsgrad mit den datenliefernden operativen Informations- und Kommunikationssystemen erreicht werden. Für die Realisierung effektiver EIS muß daher, neben der betriebswirtschaftlichen Integration in das Führungssystem, die technische Integration in die informations- und kommunikationstechnische Infrastruktur geleistet werden. Diese Integration muß das firmeneigene Informationsmanagement gewährleisten.

Bei der Realisierung von EIS in verschiedenen Unternehmen (in der Regel gab es meist mehrere Anläufe) haben sich die fehlende konzeptionelle Daten- und Dokumentmodellierung und die daraus folgende Probleme mit der qualitativen und physikalischen Datenbeschaffung (d.h. Schwierigkeiten bei der Identifikation originärer Informationsquellen) als Haupthindernis gezeigt. So sind in einigen Unternehmen die gewünschten Daten aufgrund nichtoptimaler Datenstrukturen (wirtschaftlich) nicht ökonomisch extrahierbar, bzw. es existiert kein konsistenter Datenpool für diese Systeme, so daß EIS auf (teilweise willkürlich gewählten) anwendungsbezogenen Extrakten der individuellen Datenverarbei-

tung basieren muß. Dadurch sind auch einige wichtige Funktionen nicht oder nur schwer realisierbar, z.B.:

- (Exception-)Reporting, da Dokumente aus verschiedenen Quellen unterschiedlich strukturiert sind;
- Individuelle Drill-Down-Techniken, d.h. zielorientiertes Abrufen von Detailinformationen nach der vorher definierten logischen Abhängigkeit.

Die geschilderten Probleme hängen mit der häufig nicht vorhandenen Einbettung von EIS in übergreifende Konzepte der Bürokommunikation und mit dem Fehlen eines gebietsüberschreitenden Informationsmanagement zusammen.

2.5 World Wide Web / Elektronisches Publizieren

Das World-Wide Web (WWW) besteht aus vernetzten Seiten ohne erkennbare Struktur. Bei der Gliederung der Seiten ist jeder Autor völlig frei, weil sich mit der im WWW verwendeten Sprache HTML (bisher in der Regel HTML 3.0 als Standard) nur die Gestaltung des Inhalts steuern läßt, und nicht die inhaltliche Struktur eines Dokuments. Dies hat sich zwar inzwischen bei HTML 4.0 als eine Reaktion auf das Konzept von XML etwas gebessert, es ist dennoch z.B. nicht möglich, im Rahmen einer Qualitätssicherung zu überprüfen, ob eine Seite ganz bestimmte, dort notwendige Angaben enthält. Werden sich z. B. im ZT-Intranet alle laufenden Projekte auf einer HTML-Seite aufgelistet, kann nicht überprüft werden, ob dort jeweils ein Ansprechpartner mitsamt Telefonnummer und e-mail-Adresse genannt wird. Die Dublin-Core-Initiative versucht zwar dies in Griff zu bekommen, sie scheint jedoch bisher lediglich bei den Bibliothekaren auf gewisse Resonanz zu stoßen (vgl. *Dublin Core Metadata 1997*).

Wegen fehlender inhaltlicher Strukturen in den WWW-Dokumenten gibt es nur zwei Möglichkeiten, auf Informationen zugreifen zu können: durch Navigation (Browsen) und durch den Einsatz von Suchmaschinen. Bei der Navigation ist der Leser auf die Qualität der vom Autor vorgegebenen Hypertext-Struktur angewiesen, die nicht immer mit der gewünschten Zugriffsstruktur übereinstimmt. Es gibt daher Suchmaschinen im WWW, mit denen über Stichwörter oder über Kombinationen von Stichwörtern nach Dokumenten mit der gesuchten Information recherchiert werden kann. Suchmaschinen basieren auf dem Prinzip der gewichteten Volltextsuche in möglichst vielen Seiten im WWW. Es ist nicht möglich, die Suche auf bestimmte inhaltliche Strukturen einzuschränken. Ein großer Vorteil des WWW liegt im elektronischen Publizieren von Dokumenten. Durch Plug-Ins können moderne WWW-Browser inzwischen auch Dokumente mit proprietären Formaten (PDF, WinWord, usw.) darstellen. Damit lassen sich diese zwar anzeigen, lesen und drucken; eine Weiterverarbeitung, wie etwa bei einer automatischen Indexierung durch Suchmaschinen, setzt proprietäre Software zur Konversion des verwendeten Formats voraus. Zusätzlich wird durch den Einsatz von proprietären Browsern, d.h. solche, die i.a. nur bestimmte HTML-Sprachvarianten unterstützen, der weltweite Zugriff auf die gesuchte Information zunehmend eingeschränkt.

Zusammengefaßt ist das WWW ein hervorragendes Medium, Informationen zu publizieren; zur maschinellen Weiterverarbeitung oder gar Verwaltung der publizierten *Informationen* dagegen ist das WWW ungeeignet, weil es keine inhaltlichen Strukturen kennt.

2.6 Konsistenzproblematik / Dokumentenaustausch (Protokolle)

Die inhaltliche Konsistenz von Dokumentation mit inhaltlichen Bezügen kann verletzt werden, wenn Änderungen nicht adäquat propagiert werden. Es gibt hier drei Klassen von inhaltlichen Bezügen in elektronischer Dokumentation: Wiederverwendung, Querverweise und sonstige Bezüge zwischen *Informationseinheiten*.

Bei der Wiederverwendung existieren die gleichen Informationseinheiten an verschiedenen Stellen im Dokument, evtl. angepaßt an den jeweiligen Kontext. Wird nun eine dieser Stellen verändert, müssen die Änderungen an allen anderen nachgezogen werden, um die inhaltliche Konsistenz beizubehalten.

Bei Änderungen am Absprungpunkt oder am Einsprungsziel eines Querverweises sind verschiedene Folgeaktivitäten denkbar, um wieder eine inhaltlich konsistente Dokumentation zu erhalten. Entweder der Verweis ist inhaltlich obsolet oder Dokument strukturell unvollständig, dann muß er gelöscht werden. Oder es sind keine Folgeänderungen nötig und der Verweis bleibt unverändert. Zuletzt kann es sein, daß durch die anfängliche Änderungen ein umfangreiches Umstrukturieren oder Überarbeiten von Teilen der Dokumentation notwendig wird.

Weitere inhaltliche Bezüge liegen schließlich z.B. zwischen Übersetzungen vor: Nach Änderungen am Original müssen alle bereits existierenden Übersetzungen nachgezogen werden.

Mit der Konsistenzproblematik hängt auch die Frage des Dokumentenaustausches und der Protokolle zusammen. Unter einem Protokoll wird im weiteren die Gesamtheit aller für den Daten- bzw. Dokumentenaustausch über eine Schnittstelle vereinbarten Regeln verstanden. Darunter fällt auch die in der Datenübertragung als Protokoll bezeichnete Verfahrensvorschrift. Mit dieser Vorschrift werden dann sowohl die Formate der "Nachrichten" als auch der Datenfluß bei der Datenübertragung festgelegt (vgl. auch ISO-Schichtenmodell).

Der gängige, auf dem sog. Document Interchange Format (DIF; 7-bit-Codierung) oder auf der Document Content Architecture (DCA; 8-bit-Codierung) bzw. auf den Formaten RTF und MIF basierende Dokumentenaustausch führt durch die Vermischung von logischer und physischer (Layout) Struktur zu einer erheblichen Layoutveränderung bzw. zu einem Layoutverlust und somit auch

zu einer Minderung des Informationswertes bei ausgetauschten Informationseinheiten.

3. Randbedingungen

3.1 Kundenforderungen

Technische Dokumentation dient zur Kommunikation des technischen Wissens zwischen Hersteller und Kunden. Gleichzeitig sind die in den Dokumenten enthaltenen Informationen für den Betreiber von langlebigen technischen Systemen wertvoll und bedeutsam, weil zur Wartung und Reparatur auf die benötigten Unterlagen auch noch nach Jahren oder Jahrzehnten zugegriffen werden muß. Angesichts zunehmend kürzerer Innovationszyklen in der Softwareentwicklung müssen daher bei der Archivierung von elektronischer Dokumentation auch die zur Ansicht und Weiterverarbeitung benötigten Applikationen aufgehoben werden.

Dies wurde bereits in den 80er Jahren vom amerikanischen Verteidigungsministerium (DoD) erkannt. Um die allein durch den Besitz eines komplexen technischen Systems anfallenden Kosten ('cost of ownership') zu senken, forderte das DoD im Rahmen der CALS-Initiative u.a. von seinen Zulieferern, die gesamte technische Dokumentation in einem herstellerunabhängigen Format ('CALS-DTD') auszuliefern, mit dem auch gleichzeitig die Inhalte der Dokumentation explizit gekennzeichnet wurden. Grundlage dieses Formats ist der ISO-Standard SGML, der seitdem immer häufiger zur Definition von branchenspezifischen Austauschformaten herangezogen wird.

Im folgenden werden vier wichtige Initiativen aus der unmittelbaren Vergangenheit kurz beschrieben:

- **TBL-P Projekt**

Die Deutsche Telekom fordert im Rahmen des TBL-P Projekts, an dem u.a. Alcatel SEL und Siemens ÖN als Zulieferer beteiligt sind, die technische Dokumentation in SGML bezüglich einer vorgegebenen, inhaltlichen Struktur (*Telekom DTD*), in der Informationseinheiten wie z.B. *Kommando*, *Warnungshinweis* usw. gekennzeichnet sind.

- **European Forum for Telecom Industry Information Interchange (EFTI3)**

Aus dem ESPRIT-Projekt *TecPad* (EP 8793) hervorgegangen ist die Initiative *EFTI3*, an der neben den Herstellern Siemens ÖN, Italtel, Ericsson, Alcatel und Nokia auch die Betreibergesellschaften Teledanmark und British Telecom beteiligt sind. Dieses Forum hat inzwischen eine Austauschstruktur verabschiedet, die es einem Kunden ermöglicht, die Dokumentation von verschiedenen Herstellern mit einer stets gleichen Struktur zu beziehen um sie danach intern weiterzuverarbeiten.

- **Telecommunications Industry Forum (TCIF)**

In Amerika gibt es mit TCIF eine ähnliche Initiative wie *EFT13*, an der u.a. Bellcore, CISCO Systems, Lucent Techn., Newbridge, Northern Telecom und Siemens Stromberg-Carlson auf der Herstellerseite beteiligt sind sowie u.a. AT&T, Bell Atlantic, MCI Communications, Pacific Bell und Sprint Corp. auf der Betreiberseite.

Auch in der Automobilbranche wird zunehmend SGML zum Austausch von inhaltlich strukturierter technischer Dokumentation eingesetzt. In dem sehr ambitionierten Projekt *MSR* (Messen, Steuern, Regeln) z.B. sind Mercedes Benz und Bosch federführend beteiligt, zudem Volkswagen, BMW und Porsche.

- **SAEJ2008**

In Amerika gibt es den Industriestandard SAE J2008 zum Austausch von strukturierter Servicedokumentation für alle umweltrelevanten Bestandteile von Automobilen. Vom California Air Regulations Board (CARB) wird inzwischen gefordert, daß der Austausch von Servicedokumentation konform zur SAE J2008 stattfinden muß, um in Kalifornien ab dem Jahr 2002 Fahrzeuge verkaufen zu können.

3.2 Qualitätsmanagement durch ISO 9000

Die Dokumentation von Verfahrensanweisungen, Ergebnissen und Produkten bilden einen zentralen Aspekt beim Qualitätsmanagement. Darüber hinaus spielen Dokumente eine tragende Rolle im Rahmen der Produkthaftung.

Dies alles findet in den europäischen Normen EN ISO-900x seinen Niederschlag, schwerpunktmäßig im Kapitel 4 *Forderung an die Qualitätssicherung / Qualitätsdarlegung* und speziell in den Unterkapiteln 4.4 *Designlenkung*, 4.5 *(Lenkung der Dokumente)* und 4.16 *(Lenkung von Qualitätsaufzeichnungen)*.

Auch wenn die EN ISO 900x keine direkten Vorgaben für Art, Ausführung und Umfang von Dokumenten gibt, so stellt sie aus qualitativer Sicht hohe Ansprüche an Dokumentation als solche, vor allem an das Dokumentenmanagement. Es wird gefordert, daß Dokumente mit korrektem Ausgabestand und richtigen Inhalt zum richtigen Zeitpunkt am richtigen Ort sind. Ungültige und/oder überholte Dokumente sind sofort an allen Stellen zu entfernen. Dokumente, die aus rechtlicher Sicht oder zur Erhaltung des Wissensstandes aufbewahrt werden, sind entsprechend zu kennzeichnen. Änderungen von Dokumenten und Daten sind, wo durchführbar, erkennbar in der Dokumentation selbst oder in geeigneten Anlagen ausgewiesen werden.

Für Designunterlagen gilt insbesondere, daß das Designergebnis in einer Form dokumentiert werden muß, das in Bezug auf die Designvorgaben verifiziert und validiert werden kann. Hiermit wird grundsätzlich eine Beziehungstransparenz

gefordert, welche das State-of-the-art Dokumentenmanagement a priori nicht erfüllen kann.

4. Optimierungsbedarf bei der Dokumentation

4.1 Inhaltsorientiertes Strukturieren von Dokumenten und Information

Dokumente, präziser dokumentgebundene Informationen, spielen in allen Wertschöpfungsketten des Hauses eine zentrale Rolle. Sie dienen der Sicherung und Weitergabe von Ergebnissen und Know-how für Produkte und Prozesse. Sie sind Grundlage für Vertrieb und Marketing, Entwicklung und Betrieb, Service und Wartung.

Auf Grund dieser enormen Bedeutung von Dokumentation im Wertschöpfungsprozeß, den stetig wachsenden Kundenanforderungen, den Erfordernissen in Bezug auf die Informationsdistribution und das Qualitätsmanagement ist es notwendig, auch bei Dokumenten die Informationssicht in den Vordergrund zu stellen. Basis hierfür ist eine logisch inhaltliche Strukturierung von Information und Dokumentation, die erforderliche Transparenz und Präzision in allen Phasen beim Umgang mit Information und Dokumentation sicherstellt. Bei der dazu notwendigen technologischen Grundlagen für modulares Design und Reuse von Dokumentation und Information stehen

- die Trennung von logischem Inhalt und grafischer Darstellung
 - die Modellierung komplexer Dokument- und Informationsstrukturen
 - und Konzepte und Methoden der Wiederverwendung
- im Mittelpunkt.

Eine Generalisierung über jegliche Art von Dokumentation und Information kann aus dem bisher gesagten nicht abgeleitet werden. Die folgende Auflistung zeigt daher exemplarisch, in welchen Bereichen potentiell Handlungsbedarf besteht.

Technische Unterlagen

- Produktunterlagen
- Angebotsdokumentation
- Lasten-, Pflichtenhefte (bzw. A10, A20, A30)
- Entwicklungsdokumentation, Software- und Hardwareunterlagen
- Bedienhandbücher, Anleitungen
- Service- und Wartungsunterlagen

Prozeß- und projektrelevante Unterlagen

- Projektsteuerung
- Projektunterlagen
- QM-Unterlagen
- E-mail
- Standardisierter elektronischer Briefverkehr

Unternehmensspezifische Informationen

- Informationsbasen zur Know-how-Sicherung
- Betriebswirtschaftliche Informationen und Unterlagen
- Betriebsführungsrelevante Informationen und Unterlagen

4.2 Konsistenzproblematik

Inkonsistenzen entstehen, wenn nach bestimmten Veränderungen am Dokumenteninhalte die dadurch notwendig gewordenen Folgeänderungen unterbleiben (s.o. Abschnitt 2.6). Es gibt bislang kein Werkzeug, um die Autoren beim Erstellen einer inhaltlich konsistenten Dokumentation zu unterstützen, weil mit heutigen Textverarbeitungssystemen i.d.R. nicht der Inhalt von Dokumenten manipuliert werden kann, sondern nur die Darstellung dieser Inhalte. Lassen sich die Inhalte von Dokumenten maschinell verarbeiten, wurden also die *Information* in den Dokumenten kodiert, können die Autoren von Dokumentation jedoch hinsichtlich der oben (im Abschnitt 2.6) beschriebenen Konsistenzproblematik durch semi-automatische Verfahren unterstützt werden.

4.3 Auffinden von Information

Dokumentenarchitekturen, welche die logische und physische (Layout) Struktur von Dokumenten unabhängig von spezieller Hard- und Software (meist in hierarchischer Form) beschreiben, ermöglichen eine effektive und effiziente inhaltliche Erschließung von Dokumenten und somit auch ein verlust- und ballastarmes Retrieval von Information. Ein Vertreter solcher Architekturen ist ODA. Mit SGML wird mit einem völlig anderen Ansatz das gleiche Ziel, die Trennung von Logik und Layout, verfolgt.

In der Vergangenheit wurde dies schon bei einigen Ansätzen probiert, hierzu einige Beispiele::

- Strukturierung der Informationseinheit (Dokument) durch Aspektierung bei Siemens-Projekt CONDOR (es wurde modifiziert und wesentlich erweitert von GOLEM-System übernommen).
- Textfacetten, d.h. ein Objekt ist nicht als ganzes beschrieben, sondern nach seinen Eigenschaften (z.B. Gift: Chemische Substanz, sinnliche Eigenschaften - wie z.B. Farbe, Geruch, Zustand etc. - Vergiftungssymptome, Gegenmaßnahmen etc.); leider existieren bisher keine Werkzeuge, die diese Beschreibung zumindest semi-automatisch ermöglichen.

Der Benutzer konnte dann durch Angabe von Aspekten und/oder Teilschemata/Textfacetten festlegen, welche Textteile für die folgende Recherche berücksichtigt werden sollen. Ähnlich kann man dann auch bei den Dokumentensammlungen vorgehen, die schon mit Werkzeugen erstellt wurden, die eine Trennung von Inhalt und Layout ab Erstellung ermöglichen. Einerseits werden die Pro-

bleme mit der Indexierung von großen Dokumenten in die inhaltliche Beschreibung ihrer thematisch zusammenhängenden Komponenten verlagert, andererseits werden relevante Informationen in einer ausreichenden Qualität und Quantität (ohne den Ballast des sonstigen informationslosen Dokumententextes) schnelles und präziser ermittelt sowie an den richtigen Ort verteilt (Prämissen der Informationslogistik).

Die verbindliche Verwendung von geeigneten, einheitlichen Werkzeugen zur inhaltlichen (logischen) und physischen (Layout) Strukturierung von Dokumenten kann auch den o.g. Engpaß (d.h. die fehlende konzeptionelle Datenmodellierung und daraus folgende Probleme mit der Identifikation originärer Informationsquellen) bei dem Einsatz von Führungsinformationssystemen (EIS) beseitigen helfen.

4.4 Durchgängigkeit im Prozeß durch Dokumentenaustausch

Dokumente werden oft benötigt, um die Informationen zwischen den Beteiligten eines Arbeitsprozesses auszutauschen. Dies trifft auf Verwaltungsabläufe ebenso zu, wie auf die Weitergabe von Entwicklungsunterlagen an die Autoren von Servicedokumentation oder die Weitergabe der Daten in einem CIM-Prozeß (CAD/CAP/CAM - Kette).

Das Konzept elektronischer Verfahrensketten, egal ob beim Publizieren oder in anderen Anwendungsgebieten, wie z.B. bei der o.g. computerintegrierten Fertigung CIM (CAD/CAP/ CAM - Kette) oder Entwicklungsdokumentation, lebt davon, daß ein Kernbestand von Basisdaten (oder Dokumente bzw. ihrer Teile) weitergereicht und in den nächsten Stufen des Prozesses angereichert wird (Wiederverwendung von Dokumentation oder ihrer Teile und nutzungsgerechte Informationsverteilung). Dokumente als Container für Informationen ermöglichen jedoch keinen durchgängigen Informationsfluß zwischen den beteiligten Funktionen, wenn mit dem Austausch auch ein Konvertierungsaufwand verbunden ist oder sogar ein erneutes Eingeben.

Bei Dokumentformaten, mit denen nicht der Inhalt, sondern die Darstellungsanweisungen kodiert werden, kann kein Informationsaustausch erfolgen, lediglich ein Layout-Austausch. Damit ist es jedoch nicht möglich, die in einem Dokument enthaltenen Informationen wiederzuverwenden, automatisch weiterzuverarbeiten oder für ein nutzergerechtes Verteilen aufzubereiten.

Um beim Dokumentenaustausch einen durchgängigen Informationsfluß zu ermöglichen, bedarf es einerseits eines gemeinsamen Verständnisses der auszutauschenden Information sowie eines Austauschformats mit dem sich die Informationseinheiten im Dokument beschreiben lassen.

4.5 Automatisiertes Erzeugen von Dokumentation

Dokumentation wird bislang fast ausschließlich von Redakteuren geschrieben. Wenn sich Dokumentation auch nicht vollständig automatisiert erzeugen läßt, kann der Autor jedoch weit mehr Unterstützung durch das Textverarbeitungssystem erwarten, als z.B. das Erstellen von Verzeichnissen oder das automatische Numerieren von Überschriften. Den Inhalt der Dokumentation bereitzustellen, bleibt Aufgabe des Redakteurs; er kann jedoch sowohl beim Erzeugen der erforderlichen Dokumentstruktur und beim Auffüllen dieser Struktur mit bereits existierenden Informationen unterstützt werden, als auch bei einem späteren Veredeln der Dokumentation (z.B. mit Verweisen auf andere Quellen, Indizes und Verzeichnissen).

Steht fest, welche Struktur das zu erzeugende Dokument haben muß, welche Informationen möglich und welche notwendig sind, kann eine leere Struktur des Dokuments automatisch erzeugt werden, die der Autor oder ein Generator nur noch mit den jeweiligen Inhalten füllen muß. Die leere Struktur selbst kann auch aus den Informationen in anderen Dokumenten gewonnen werden. Beispielsweise kann die Struktur einer Programmbeschreibung aus den Strukturen des Programms erzeugt werden. Auch lassen sich kundenspezifische Varianten aus einem umfassenden Dokumentationsbestand generieren oder vorliegende Bestandteile automatisch in einem Dokument integrieren. Grundvoraussetzung ist, daß die Struktur des zu erzeugenden Dokuments explizit ist. Gibt es bereits Ausgangsmaterial für das zu erzeugende Dokument, muß spezifiziert sein, welche Informationen an welcher Stelle und in welcher Form im Zieldokument übernommen werden sollen.

Automatisiertes Erzeugen von Dokumentation ist vor allem aus Kosten- und aus Qualitätssicht bedeutsam. Aus Kostensicht, weil Dokumentation gezwungener Maßen meist als "Abfallprodukt" gesehen wird. Aus Qualitätssicht, weil nur eine weitgehend automatisierte Verarbeitung die erforderliche Qualität sichert. Es kann davon ausgegangen werden, daß die hier betrachteten Information und Dokumentation i.a. logisch inhaltlich gut bis sehr gut strukturierbar ist. Es liegt also die Forderung nach einer automatisierten Erzeugung nahe.

Ein zentraler Schwerpunkt dabei ist die strukturelle Transformation von Dokumenten (Werkzeuge Metamorphosis, Balise, etc.). Damit ist es möglich, Dokumentation und Information für die jeweiligen Bedürfnisse maßzuschneidern. Aus der Distributionssicht führt dies zu *Single-Source-Publishing*, aus konzeptueller Sicht zu *Single-Source-Information*.

4.6 Single Source Publishing /Single Source Information

Technische Dokumentation wird i.d.R. in mehreren Varianten publiziert, entweder um bestimmte Zielgruppen (Sprachvarianten, nutzungsspezifische Varianten, ...) oder um verschiedene Publikationsmedien (HTML, PDF, Papier, ...) zu

bedienen. Bislang wird zunächst ein Originaldokument erstellt, aus dem anschließend alle benötigten Varianten erzeugt werden. Dabei besteht zum einen die Gefahr, daß es bei Änderungen an redundant dokumentierten Informationen zu inhaltlichen Inkonsistenzen kommen kann, wie bereits im Abschnitt 'Konsistenzproblematik' besprochen wurde. Zum anderen ist beim Erzeugen gerade von verschiedenen Publikationsvarianten oft ein manueller Konvertierungsaufwand erforderlich, der zudem bei jeder nachfolgenden Veränderung erneut anfällt.

Hier sollte die Information daher redundanzfrei in einem einzigen Dokument verwaltet werden, aus dem - nach Bedarf - das jeweils gewünschte Publikationsformat generiert wird (*Single Source Publishing*). Die Voraussetzung dafür ist erneut, daß die Informationen im Dokument zugreifbar sind und daß spezifiziert ist, wie diese Informationen in den verschiedenen Formaten dargestellt werden soll und welche Sichten auf das Dokument jeweils gewünscht werden..

Die Vorteile eines solchen Ansatzes sind:

- Ändern sich die Anforderungen an das Layout, muß das Ausgangsdokument nicht verändert werden, sondern lediglich die Anweisung, wie die Informationseinheiten im jeweiligen Format dargestellt werden sollen. Beispielsweise wird die im WWW verwendete Sprache HTML zur Darstellung von Dokumenten oft für neue Versionen des Browsers erweitert. Damit ältere Versionen nicht antiquiert wirken, muß lediglich der Einsatz der neuen Sprachmittel neu spezifiziert werden; das Generieren der Publikation erfolgt automatisch.
- Aus einer Quelle können über formatspezifische Filter die Publikationen für verschiedene Medien bedient werden, ohne manuellen Zusatzaufwand.

Die gewünschten Varianten können als unterschiedliche Sichten aus einem redundanzfreien Dokumentationsbestand generiert werden; die Pflege der Informationen erfolgt zentral.

Qualitativ hochwertige Dokumentation und Information setzt auch korrekte, widerspruchsfreie Primärinformation voraus. Unter Abwägung von Aufwand und Nutzen ist von einem Idealszenario auszugehen, bei dem die Quelle für jedes Dokumentationsfragment und jede Informationseinheit eindeutig verifizierbar ist. Quelle bedeutet in diesem Zusammenhang nicht ausschließlich ein technisches System, wie z.B. eine Datenbank. Quelle meint ebenso den menschlichen Autor. Dies gestattet, sowohl einzelne Dokumente als auch ganze Dokumentverbände als Informationsressource, ähnlich einer Datenbank zu nutzen. Voraussetzung hierfür ist

- die Einführung konzeptuellen Dokument- und Informationsschemata (Document Type Definition - DTD);
- die formale inhaltlich, logische Strukturierung (Logical Markup).

Die Datenintegrität von Datenbank-Management-Systemen kann verständlicherweise nicht übertragen werden.

5. Weltweite Aktivitäten in Auswahl

- Bei der Firma Ericsson läuft seit 1991 ein strategisches Programm zur Geschäftsprozeß-Optimierung mittels durchgängiger, SGML-basierter Dokumentation, mit einem Aufwand von ca. 140 Mio. DM (bis Ende 1995). Dazu folgendes Zitat aus: *A Guide to SGML at Ericsson*, Spring 1995: "By being the first to deliver documentation using electronic media that adheres to the SGML standard, Ericsson can influence or even define the branch standard for this application, and side-step the costly and time-consuming process of converting our documentation".
- Der Deutsche Telekom fordert im Rahmen des TBL-P Projekts, an dem u.a. Alcatel SEL und Siemens ÖN als Zulieferer beteiligt sind, die technische Dokumentation in SGML bezüglich einer vorgegebenen, inhaltlichen Struktur (Telekom DTD), in der Informationseinheiten, wie z.Z. Kommando, Warnungshinweis u.s.w., gekennzeichnet sind.
- Bei der Firma Newbridge wird die gesamte technische Dokumentation auf der Basis von SGML erstellt, verwaltet, gepflegt und publiziert. Zitat (<http://www.newbridge.com/Information/CustDoc/index.html>): "Newbridge has adopted SGML, an ISO standard for the markup of text and definition of document hierarchy, as the basis for its development of online documentation. This standard allows Newbridge not only to create a better, more sophisticated product, but also to deliver it in ways that make sense for customers' specific needs. In the future, Newbridge customers can choose to receive their documentation on CD-ROM, paper or across the Internet."
- Siemens ÖN Breitbandnetze hat eine Allianz mit Newbridge (Kanada) gebildet. Zu diesem Zweck wird die Kundendokumentation bestimmter Produkte in SGML ausgetauscht. Damit Siemens und Newbridge (NB) diese Dokumentation jeweils bei sich in einen Produktionsprozeß einfließen können, wird Siemens-Dokumentation auf NB-DTD transformiert und umgekehrt (man spricht dabei auch über Allianz-Dokumentation).
- British Telecom betreibt eine zentrale Anlieferungsstelle, bei der die technische Dokumentation in SGML vom Hersteller empfangen, kontrolliert, freigegeben und als SGML selber oder mittels einer Echtzeit-Konvertierung nach HTML im BT-Intranet verteilt wird.
- Verschiedene Fachverlage setzen auf SGML als internes Verwaltungsformat für seine Publikationen auf. So z.B. nutzen die Verlage Dr. Otto Schmidt und Carl Heymanns SGML zur Qualitätssicherung ihrer Erzeugnisse.
- Beim Deutschen Institut für Normung (DIN) gibt es ein Projekt, die Normen mit SGML inhaltlich zu strukturieren. Mit der vom Beuth Verlag im April 1997 publizierten CD-ROM mit SGML-Daten können Firmen die für sie relevanten

Normen durch ein automatisches Konvertieren nach HTML in ihrem Intranet publizieren.

- In der zivilen und militärischen Flugzeugindustrie (Airbus, Lufthansa, British Aerospace, Royal Air Force, ...) wird die technische Dokumentation mit SGML ausgetauscht und von den Fluglinien automatisch weiterverarbeitet und publiziert.
- Die Firma Novell erzeugt, verwaltet und publiziert ihre Online-Dokumentation auf der Basis von SGML.
- Seit zwei Jahren gibt es die - inzwischen unter der Schirmherrschaft des *World-Wide Web Consortiums (W3C)* stehende - Initiative *Extensible Markup Language (XML)*, mit der sich im Internet zukünftig Informationselemente publizieren lassen, und nicht mehr nur darstellungsorientierte HTML-Seiten. Inzwischen sind Microsoft und Netscape an der Spezifikation von XML beteiligt. Microsoft hat zudem angekündigt, daß XML "im Internet Explorer 4.0 verwendet werden wird".

6. Resume und zugleich Zusammenfassung

Durch die Trennung der inhaltliche und logischen Struktur wird die in den Dokumenten enthaltene Information explizit und maschinell weiterverarbeitbar gemacht. Durch diese Eigenschaften können dann die Grundprämissen der Informationslogistik erfüllt werden. Dabei werden sowohl strukturierte Dokumente als auch Datenbanken als Container für Informationen bzw. Informationseinheiten (Wissenseinheiten) angesehen, die editiert und gepflegt werden sowie recherchierbar und adäquat publizierbar gemacht werden müssen.

Literatur:

Hinweise zu Aktivitäten im Abschnitt 5 und zu Initiativen im Abschnitt 3.1 können alle im Internet gefunden werden. Die im Text erwähnten ISO-Normen sind allgemein bekannt. Zur effektiven Nutzung von strukturierten Dokumentation in verschiedenen Anwendungen (wie z.B. in Information-Retrieval-Systemen oder in Führungsinformationssystemen) findet man in der Literatur bisher so gut wie keine Hinweise (vgl. die im Literaturverzeichnis angegebene Conference Proceedings).

Dublin Core Metadata: <URL:http://purl.org/metadata/dublin_core> 1997.

Heeter, G.: Technical Information - the North American Customer. In: Proceedings der 7. Fachtagung Technische Dokumentation, Siemens AG, München 1998.

keine Hinweise (vgl. die im Literaturverzeichnis angegebene Conference Proceedings).

Dublin Core Metadata: <URL:http://purl.org/metadata/dublin_core> 1997.

Heeter, G.: Technical Information - the North American Customer. In: Proceedings der 7. Fachtagung Technische Dokumentation, Siemens AG, München 1998.

Möller, H.: Konsistentes Dokumentenengineering mit protokollierten Dokumenten. Diss. Berlin: FU Berlin, FB Mathematik und Informatik 1997.

SGML Europe '96: Conference Proceedings (Conf. Chair: Pamela Gennusa), München 1996. Alexandria (USA): Graphic Communication Ass. 1996.

SGML / XML '97: Conference Proceedings, Washington D.C. 1997. Alexandria (USA): GCA Graphic Communication Ass. 1997.

SGML / XML Europe '98: Conference Proceedings, Paris 1998. Alexandria (USA): GCA Graphic Communication Ass. 1998.

Library within Cyberspace Paradigm: the Integrated Virtual Library Concept

Saša Skenderija

Institute of Information Studies and Librarianship, Charles University Prague
The State Technical Library, Prague
EMAIL: skenders@cuni.cz
URL: <http://www.cuni.cz/~skenders>

The role of libraries, librarians and information specialists in the situation when libraries and traditional information institutions are not sole and exclusive provider of relevant information (potentially useful to users) is one of the most frequently discussed questions of our day.

At present, when cyberspace becomes a global communication paradigm of its kind, a traditional "institutional" approach to problems of establishment and management of public information structures is often regarded as unsuitable. The new communication environment emerging in cyberspace is highlighted by open, decentralised and flexible "hyperstructures" as well as by interactive client/server procedures. This constitutes an exact opposite to traditional centralized and institutionally managed library and information structures and processes. However, it is apparent that in practice some of the centralized or institutionally managed library and information procedures (e.g. shared cataloguing, co-ordinated acquisition, union catalogues, interlibrary loan services etc.) are irreplaceable, for both economic and procedural reasons.

The Integrated Virtual Library concept (abbreviated as **INVIK** - the abbreviation stands for "integrovaná virtuální knihovna" or "Integrated Virtual Library" in English) presents an alternative approach to questions of role and function of the modern library conditioned by an interactive decentralised and open communication structures and processes which is marked and influenced by medial properties of cyberspace. This concept was aimed at the establishment of the fundamentals for **functional co-existence of the traditional library with the "virtual library"** in the sense of integrated information service which is able to combine electronic access to both

"material" and "virtual" sources of information with a wide variety of methods of delivery and mediation. The scope and quality of services offered is based both on the ability to implement new network sources, technologies and co-operative models into outputs of traditional services as well as on the ability to merge proven traditional library methods and techniques (know-how) in conditions of the newly established global network environment. The desired result is an amalgamation of all appropriate procedural models and techniques, of all accessible sources and means which would provide comfortable and barrier-free access to relevant information sources.

The Integrated Virtual Library - INVIK presents a concept, an idea of communication system (i.e. not only technology in the sense of hardware and software) integrating various sources and services on one interactive communication channel in a medial environment of WWW; i. e. the sources and services that are separated from each other in space and time.

The INVIK system is functioning in the decentralised, pluralistic and open environment of cyberspace. The effectiveness of an integrated virtual library as integral and flexible library/information service depends also on an optimal interconnection and co-ordination with a local and global library and information environment. The aim of such an interconnection is to obtain a union of all sources, methods, means and techniques available to respond to user needs as fast, cheaply and comfortably as possible.

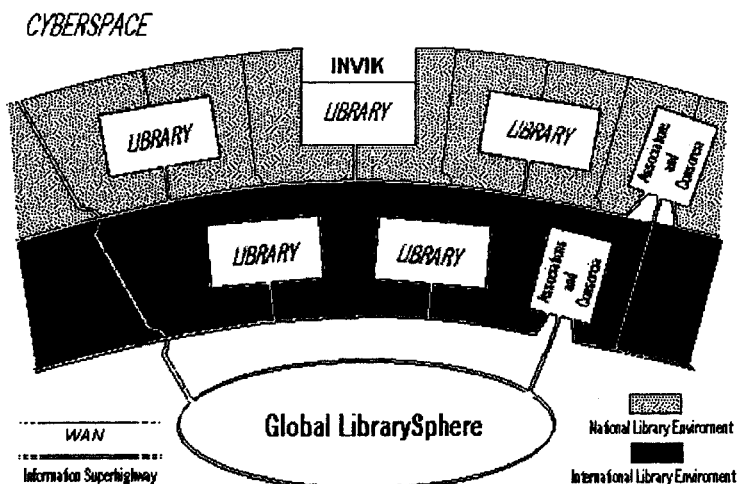


Fig. 1: Integrated Infosphere

The scheme enclosed (see Fig.1) shows the position of the Integrated Virtual Library - INVIK within the newly emerging global library and information

environment - the *integrated infosphere*. It also demonstrates the obsolescence and inadequacy of the concept of centralized library and information structures as the only basis for a solution of present global information problems. The fundamental deficiency of all centralized library and information structures stems from the fact that inevitably they are centered around institutions. On the contrary, in cyberspace there is no fixed center. It is substituted by "variable centers" represented by **individual users at the moment of placing their information request**.

The homogenous centralized library information structures and bodies together with their established co-operative procedures and models (union catalogues, overall digitising of traditional sources etc.) admittedly improve the logistic bases of the integrated virtual library but they are not the only basis nor the only possibility for effective co-operation and sharing of procedures in library and information environment. Non-existent or defunct union catalogues may be to some extent replaced by decentralised search procedures in WWW OPACs of relevant libraries or by application of network search/retrieve protocols like Z39.50. Similarly, slow and costly processes of "total digitising" of traditional holdings may be replaced by local *"just-for-you-digitising"* (*"digitising-on-demand"*) procedures etc.

For the sake of the best service to its user, INVIK is open to all forms and means of purposeful co-operation with local, international and global library and information institutions, structures and bodies. The concept of INVIK aims at supporting the establishment of a co-operative library information environment which is open, pluralistic, flexible and decentralised and where the quality of user-oriented services provided by any co-operating library is based on **an efficient utilisation of the most powerful procedures available in participating libraries**.

The INVIK concept is a conceptual and methodological basis of the project **INVIK STL – Integrated Virtual Library of the State Technical Library in Prague** <<http://www.stk.cz/en/>>. On March 7th 1998, STL celebrated the 1st anniversary of launching INVIK for public use. The users reaction is quite convincing - between October and December 1997, INVIK was accessed more than 25.500 times. At the same time, the INVIK concept is becoming an alternative platform for interlibrary cooperation among Czech research and academic libraries.

SUPPORT OF THE INFORMATION NEEDS OF SCIENTISTS AND THE CHANGING PATTERN OF SCHOLARLY COMMUNICATION

Vladimir T. Borovansky

University Libraries
Arizona State University
Tempe, AZ 85287-1006 USA
iacvtb@asuvm.inre.asu.edu

"The delivery of high quality information services is the core mission of science and engineering libraries".

This paper will address the changing pattern of scholarly communication along with its impact on academic libraries and how it is influenced by electronic information. In addition, budgeting changes in this new environment, the need for improved information literacy and restructuring academic libraries will also be addressed. Also possible changes in scholarly publishing, the persistence of paper usage in the era of electronic information and finally, the use of faculty profiles in the improvement of information flow.

We are in a transmission mode in the information age. The impact of new technologies on patterns of scholarly communication is inevitable. We can guess what it might portend for the role and the services of the library.

Until very recently collection development librarians were concerned with the completeness in the collections of a library to serve scholars' needs, who wanted it for effectiveness and efficiency. Scholars, in the sciences in particular, are frequently concerned with access to information that may affect the experiment they are conducting in their laboratory, they need access to information as soon as possible. Now completeness of collections of a new type is available. Networks give scholars access to the information stored at different locations.

This new paradigm shift also brings with it new challenges and problems. Until very recently academic libraries budgeting practices have been based on buildings, books, materials and other resources with long, useful lives. Some of these elements will be still with us, eg. the ever increasing cost of serial literature in the sciences. Now, the technological changes are introducing new problems,

rapid obsolescence and shorter useful lives, increased maintenance costs and additional capital along with copyright challenges, licensing agreements, cooperative collection development and resource sharing which means reexamination of some of the old assumptions in light of the new environment.

The availability of work stations in faculty offices and even at home and in student dormitories makes it possible for users to receive needed information without even going to the library. Also it signifies a rise in the end-user searching. This brings with it an additional need for instruction, not only when A&I databases are searched, where most end users, especially students, lack necessary expertise and text decoding skills. In addition, in today's array of new electronic resources, the users feel frequently lost. They need a librarian/information specialist as a consultant to explain these new tools.

Another element in the delivery of high quality information services to the academic institution's clientele, in this case in particular faculty, is creation of a database of "faculty profiles". ASU took the first steps in this direction in the early seventies with the first use of a database management system in 1981. Later this effort was taken over by a committee composed of representatives of the Libraries Collection Development Office, the VicePresident for Research Office and Graduate College and published first in print. Latest development has ASU subscribing to the Community of Science database and contributing to it our own faculty profiles.

Creating an environment and structure for increasing awareness of the information needs of scientists in academic institutions and accessibility of information will be perceived as a measurement of the quality and effectiveness of scholarly communication on campus.

Electronic Library "Czech Parliament"

Karel Sosna

**PhDr. Karel Sosna
Parliament of the Czech Republic
Chamber of Deputies
Parliamentary Library
Sněmovní 4
118 26 Praha 1
e-mail: sosna@psp.cz**

The project of an electronic library is also a significant step towards computerization of services of the Parliamentary Library. This library shall contain complete texts of parliamentary prints (i.e. bills, including explanations and decisions), shorthand minutes (i.e. verbatim minutes of all sessions and all voting) and other parliamentary documents from 1861 until now in an electronic form. It would contain tens of thousands of pages of text which could create a good library with hundreds of volumes.

The Library has already taken some preliminary steps in close cooperation with the Department of Information Technology of the Committee of the House of Parliament. On the basis of a proposal by a private firm ESU Praha in June 1995, the library and the Department of Information Technology submitted a proposal for an electronic library called "Czech Parliament" to the Secretary General of the Office of the House of Parliament. The Secretary general of the Office approved the project and appointed a committee of executive officers from the Office under his leadership for final consideration of the proposal before a selection procedure is announced. The committee was to consider the usefulness of the project, set forth its content and scope and try to find some means in the budget for 1996 to start the project (the budget of the Office for 1996 was already set forth at the time of approval of the project). The committee decided to separate conversion of parliamentary texts into an electronic form from production and distribution of a CD-ROM which the project counted on as a carrier of data, and to hold a selection procedure only for the conversion. At the end of 1995, a survey of prices on the market was conducted which helped us calculate the expected costs. We came to the conclusion that it is possible to carry out activities amounting to 1,000,000 CZK in 1996.

The objectives of the project are:

- To provide access to information on current and past activities of the Czech Parliament to the public, as is customary in democratic countries. At present, all parliamentary documents (prints, shorthand minutes) are stored in written form only in the Parliamentary Library. This situation is in contrast with the fact that the public is increasingly interested in activities of the Parliament. The library, with limited access for the public, may not satisfy the demand through traditional means (providing copies).
- To provide access to current parliamentary documents with regard to the application of modern technology and methods of information work,
- To preserve the "Golden Collection" of the Parliamentary Library for the future in electronic form, because the paper is slowly decomposing and it is not even possible to make copies of some originals; electronic text on a CD-ROM would mean survival of this legacy of the Parliament for another 80 - 100 years.
- To become one of the first parliaments in Central and South-Eastern Europe to create a such electronic library.

The explanation of the project furthermore argues that the investment return from the texts of the electronic library may not be adequately valued in money terms, but that its main purposes are:

- (1) To increase the prestige of the Parliament, whose activities are still hidden to the public or interpreted (filtered) through mass media.
- (2) To protect against destruction or loss. In case of a possible sudden natural catastrophe, fire, war, and subsequent destruction of the library's collections and archives, the parliamentary documents will be still available for coming generations on a substitute medium in many places, regardless of the original place of their creation and storage (it also applies in case of a destruction of the printed documents as a consequence of natural wear of unavoidable physical and chemical processes).
- (3) To increase the speed and convenience of access for the Members of Parliament and Senators to parliamentary documents as a source of knowledge and inspiration of their legislative activities.

The whole project was planned to be implemented over three years. A decision was made that the so-called "nest-like" method shall be used, i.e. the parliamentary documents will be chronologically converted into an electronic form gradually from key historical periods of our modern history in which there is the greatest interest. In 1996 the following periods of interest were covered by

document conversion to electronic format: parliamentary documents from 1989 until now (Federal Assembly, Czech National Council, House of Parliament of the Czech Republic), and from 1968 until 1970, (National Assembly, establishment of the Federal Assembly, establishment of the Czech National Council). In 1997, parliamentary documents from 1945 to 1948 (Legislative Assembly, National Assembly), 1913 (end of the Regional Congress) and 1861 (renewal of activities of the Regional Congress) are being converted into electronic form. The remaining documents shall be converted in 1998. After these activities are finished the texts shall be stored on several CD-ROMs for the purposes of the Parliament of the Czech Republic. Another selection procedure shall be held for production and distribution of the CD-ROMs.

Two private companies won the selection procedure held at the beginning of 1996. They offered the best technical and financial parameters and at the same time carried out the most successful test of conversion of 10,000 pages of text. At present, more than a half of the project has been completed and the work continues. Existing results of the project are available on Internet at the following address - <http://www.psp.cz/eknih/>. The project has found a great favor both in the domestic and international media which can be also proved by a number of registered entries into the electronic library represented for instance by 1326 entries in May 1998 alone. Internationally compared, this project presents a unique scheme with a deep historical span. For most of democratic parliaments make accessible on Internet only those parliamentary documents of the last two or three terms of office. When the whole project is finished next year, the Parliamentary Library of the Czech Republic, at least its special collection of parliamentary documents, will be open 24 hours a day.

Section 4

The Internet as Knowledge Base

(Social Aspects and Impacts, Interface Problems)

Eine empirische Studie über die Auswirkungen elektronischer Marktplätze

Wolfgang Semar

Universität Konstanz
Fakultät für Mathematik und Informatik
Informationswissenschaft

D-78457 Konstanz
E-Mail: Wolfgang.Semar@uni-konstanz.de

Inhalt

- 1 Motivation
- 2 Veränderungen durch elektronische Marktplätze
 - 2.1 Wirtschaft
 - 2.2 Verwaltung
 - 2.3 Gesellschaft
- 3 Erhebungsverfahren
- 4 Ergebnisse
- 5 Ausblick
- 6 Literatur

Zusammenfassung

Die Ausbreitung des Internets und seiner Technologie führt dazu, daß sich die gewachsenen Strukturen in Wirtschaft, Verwaltung und Gesellschaft verändern werden. Ziel dieser auf eine längere Untersuchungszeit angelegten empirischen Untersuchung ist, die Strukturveränderungen, hervorgerufen durch die Etablierung elektronischer Marktplätze, zu dokumentieren und zu interpretieren. Die den Wandel beschreibenden Hypothesen stellen dabei den Ausgangspunkt dieser Studie dar. Durch die in zyklischen Abständen durchgeführten Erhebungen werden die Hypothesen auf ihre Richtigkeit überprüft. Der Grundstein zu dieser Untersuchung wurde mit einer ersten umfangreichen Befragung gelegt, darauf aufbauend werden im Abstand von 2 Jahren jeweils Folgeuntersuchungen durchgeführt.

1 Motivation

Die Ausbreitung des World Wide Web (WWW) schreitet unaufhaltsam voran. Zur Zeit verfügen rund 10 Millionen Bürger in Deutschland über die Möglichkeit, mit Hilfe eines Internet-Dienstleisters (Provider) oder Online-Dienstes wie AOL, T-Online oder Compuserve ins Internet zu gelangen, dies sind bereits 12,2 Prozent der Bundesbürger. Nach einer repräsentativen Umfrage der Nürnberger Gesellschaft für Konsumforschung (GfK) wollen sich innerhalb der kommenden 12 Monate weitere 10 Prozent der Bevölkerung Zugang zu dem neuen Medium verschaffen.

1997 hat sich in Deutschland die Zahl der beim Deutschen Network Information Center DENIC¹ registrierten 119.491 (Februar 1998) Domains mehr als verdoppelt, weltweit gibt es bereits 29.670.000 (Januar 1998) registrierte Domains², eine Verdopplung findet hier alle 6 Monate statt. Der ehemalige Postminister Wolfgang Bötsch vertritt die Meinung, daß bis zum Jahr 2000 rund 14.000 deutsche Unternehmen planen, ihre Produkte im Internet anzubieten. Das erzielbare Umsatzvolumen schätzt die EU-Kommission auf 25 Milliarden DM bis zum Jahr 2001 alleine in Deutschland und auf 170 Milliarden DM weltweit. Bei einer so schnellen Weiterentwicklung des Webs drängt sich die Frage auf: „Welche Veränderungen ergeben sich durch das Internet?“.

Sicherlich läßt sich voraussagen, daß sich die gewachsenen Strukturen in Wirtschaft, Verwaltung und Gesellschaft verändern werden. Dies trifft auch auf den Zuschnitt der Märkte selber zu. Für Kühlen steht fest, daß "durch elektronische Marktplätze die Standortfaktoren für den Erfolg von Regionen neu bestimmt" (Kühlen 1996) werden. Die Präsentation von Produkten bzw. die Bereitstellung von Informationen und Dienstleistungen geschieht dabei auf globalen bis hin zu kommunalen elektronischen Marktplätzen, die von Kühlen als die institutionelle Realisierung abstrakter elektronischer Märkte definiert werden. Elektronische Marktplätze unterstützen in ökonomischer Hinsicht alle Transaktionsformen der Präsentation, der Geschäftsprozesse und der laufenden Kundenbetreuung, dienen aber als Bürgerinformationssysteme auch den öffentlichen Informations- und Kommunikationsbedürfnissen. Mit der Etablierung elektronischer Marktplätze ist ein tiefgreifender Strukturwandel verbunden, der die Standortfaktoren für den Erfolg von Regionen neu bestimmen wird. Um die so hervorgerufenen Strukturveränderungen festzustellen, wird eine empirische Studie über einen Zeitraum von mehreren Jahren (seit 1996) durchgeführt, die sich mit den Auswirkungen von Mehrwertdiensten durch elektronische Marktplätze in einer bestimmten geographischen Region, der ca. 20.000 Einwohner zählenden Stadt Pfullendorf³ im Linzgau, beschäftigt.

¹ <http://www.denic.de>

² <http://www.nw.com/zone/WWW/report.html>

³ <http://www.pfullendorf.de>

Der Grundstein zu dieser Untersuchung wurde bereits mit einer ersten umfangreichen Befragung gelegt. 153 Unternehmen, 228 Bürger und Bürgerinnen, 136 Schüler sowie die Mitarbeiter der Verwaltung haben die Fragen, die sich zu diesem Thema stellen, beantwortet. Die Auswertung dieser Erstuntersuchung bildet die Ausgangsbasis, darauf aufbauend werden im Abstand von 2 Jahren jeweils Folgeuntersuchungen durchgeführt.

2 Veränderungen durch elektronische Marktplätze

Das Ziel dieser auf eine längere Untersuchungszeit angelegten empirischen Untersuchung ist, die Strukturveränderungen, hervorgerufen durch die Etablierung elektronischer Marktplätze, zu dokumentieren und zu interpretieren. Die Erarbeitung und Systematisierung der Bereiche, in denen ein Strukturwandel zu erwarten ist, dient dabei als Grundlage für das Design der in regelmäßigen Abständen durchzuführenden Erhebungen. Die den Wandel beschreibenden Hypothesen stellen dabei den Ausgangspunkt dieser Studie dar. Die formulierten Thesen und die zu ihrer Operationalisierung erarbeiteten Indikatoren bilden das theoretische "Gerüst", welches in den für die Befragung in Pfullendorf entworfenen Fragebögen abgebildet wurde.

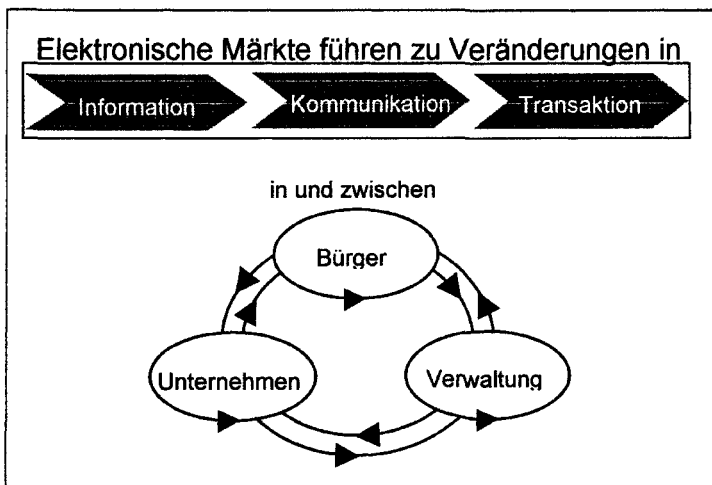


Abb. 1: Einfluß elektronischer Märkte auf das Informations-, Kommunikations- und Transaktionsverhalten von Wirtschaft, Gesellschaft und Verwaltung

Generell ist das Hauptziel regionalen bzw. kommunalen Handelns im Bereich neuer Informations- und Kommunikationstechnologien, Wettbewerbsvorteile des Standortes zu realisieren. Grabow und Korte gehen davon aus, daß "Ziele, die mit den kommunalen Aktivitäten auf dem Feld der Telematik und der Teledienstleistungen verbunden werden", sich in drei wesentliche Kategorien

einteilen lassen: "Die Verbesserung der Leistungs- und Funktionsfähigkeit der kommunalen Verwaltung [...], die Verbesserung der Wettbewerbsfähigkeit der lokalen Wirtschaft [...] und die Verbesserung der Lebensbedingungen vor Ort" (Grabow 1996). Bei der Einteilung in die Kategorien Gesellschaft, Wirtschaft und Verwaltung ist jedoch zu beachten, daß zwischen diesen Kategorien Abhängigkeiten bestehen. So werden sich Veränderungen in der Gesellschaft mit Sicherheit auf das Informations-, Kommunikations- und Transaktionsverhalten der Wirtschaft und der Verwaltung auswirken, die ihrerseits dadurch Veränderungen erfahren und durchführen werden. Ebenso werden Veränderungen in der Wirtschaft Auswirkungen auf die beiden anderen Kategorien haben. Die Frage nach der Ursache und der Wirkung ist dabei nicht immer eindeutig zu klären (siehe Abb. 1).

2.1 Wirtschaft

Die Möglichkeiten der neuen Telekommunikationsinfrastrukturen und damit die Ausbreitung von Telematikanwendungen bedeutet für die gesamte Wirtschaft eine notwendige Neuorientierung und die Redefinition von Geschäftsfeldern und -prozessen. Dies betrifft sowohl unternehmensinterne als auch die interorganisatorischen Strukturen und Abläufe und damit auch die kommerziellen und privaten Marktpartner der Unternehmen. Im einzelnen wird erwartet, daß folgende Veränderungen eintreten:

1. *Veränderung der Aufbauorganisation von Unternehmen:* Dezentralisierung der Entscheidungskompetenzen, Verflachung der Hierarchien und Verringerung der Arbeitsteilung (Rieser 1992, Tapscott 1996).
2. *Zunehmende technische Vernetzung von Unternehmen (intern + extern):* Elektronische Märkte fördern durch die Möglichkeiten, die sie für die betrieblichen Informations- und Kommunikationsprozesse bereithalten, eine technische Vernetzung (Kuhlen 1995).
3. *Veränderung der Kommunikationsbeziehungen von Unternehmen:* Bedingt durch eine verstärkte Nutzung der Mehrwertdienste des Internets, verändert sich der funktionale Kommunikationsprozess von Unternehmen.
4. *Veränderung in der Abwicklung von Geschäftsprozessen in Unternehmen:* Zur innerbetrieblichen Information und Kommunikation werden sich zunehmend die Internettechnologie und deren Mehrwertdienste durchsetzen (Palass 1997).
5. *Veränderung der Produktionsprozesse:* Einflußnahme der Kunden über das WWW auf die Wertschöpfungskette eines Unternehmens; Anpassung der Produktionsprozesse an die neuen Technologien (Bullinger 1995, Klein 1994).
6. *Enthierarchisierung der Absatz- und Beschaffungsmärkte:* Auf elektronischen Märkten werden die klassischen Funktionen des Groß- und Einzelhandels in dem Medium "Elektronischer Markt" neu abgewickelt (EM 1991).

7. *Globalisierung/Internationalisierung der Märkte und Geschäftsbeziehungen:* Die moderne Informations- und Kommunikationstechnologie fördert die Globalisierung der internationalen Wirtschaftsbeziehungen (Kumon 1993).
8. *Virtuelle Kooperationsformen von Unternehmen:* Virtuelle Unternehmen, bzw. virtuelle Unternehmensverbünde werden sich durch das Eingehen neuer Formen der Kooperation herausbilden (Kuhlen 1995).
9. *Sozioökonomischer Strukturwandel:* Verschiebung der Wirtschaftsstrukturen von der Produktionsgesellschaft zur Informationsgesellschaft (Bell 1976).
10. *Veränderung der Qualifikationsanforderungen an die Beschäftigten:* In den verschiedenen Bereichen der Informationswirtschaft entstehen neue Berufe mit neuen Qualifikationsanforderungen (Hensel 1990).
11. *Veränderungen von Arbeitsumfeldern, -bedingungen, -verhältnissen:* Durch die Möglichkeit der Informations- und Kommunikationstechnologien entstehen neue Beschäftigungsformen.

2.2 Verwaltung

Die traditionelle Versorgung der Bürger mit Information ist in der Regel nicht zielgerichtet, sondern gestreut und gleicht eher einer kommunalen Selbstdarstellung, als daß sie den eigentlichen informationellen Bedürfnissen der Bürger gerecht wird. Durch die Entwicklung der Bürgerbüros hin zu komfortableren Bürgerinformationssystemen ist es für den Bürger einfacher, die für ihn notwendige Information zu erhalten (Kubicek 1995). Dies bedingt jedoch eine mehr technik- und kommunikationsorientierte Verwaltung. Elektronische Marktplätze, die von den Verwaltungen betrieben werden, sind hierbei unter zwei Gesichtspunkten zu sehen. Einmal aus verwaltungsinterner Sicht (Horvath 1996) und zum anderen aus verwaltungsexterner Sicht. Sieht man die Verwaltung als Unternehmen an, so kann man sagen, daß die erwarteten Veränderungen aus verwaltungsinterner Sicht sich im wesentlichen mit den Veränderungen, die in der Wirtschaft erwartet werden, decken. Betrachtet man die verwaltungsexterne Seite, so lassen sich zwei wesentliche "Kunden der Verwaltung" bestimmen, die Bürger und die Wirtschaft bzw. andere Verwaltungen. Auf der externen Seite wird erwartet, daß folgende Veränderungen eintreten.

1. *Verbesserung der Kommunikation:* Unabhängige Öffnungszeiten sowie moderne Informations- und Kommunikationstechnologie fördert die Kommunikation mit dem Bürger und der Wirtschaft.
2. *Verbesserung der Informationsbeschaffung:* Zusatzinformationen durch den Einsatz von Hypertexten sowie eine anonyme Informationsbeschaffung, die dem Bürger und der Wirtschaft eine Verbesserung der Informationsbeschaffung erlaubt.

3. *Verbesserter Zugang und Abwicklung von Verwaltungsdienstleistungen:* Durch die Transaktionsmöglichkeit auf elektronische Marktplätze erfolgt eine Verbesserung in der Bearbeitung der Verwaltungsdienstleistungen.

2.3 Gesellschaft

Mit dem Einsatz und der Nutzung neuer Medien- und Kommunikationssysteme wie Internet verlagert sich die Kundenschnittstelle in den privaten Haushalt (Bievert 1994). Das Ergebnis dieser Verlagerung ist, daß private Haushalte nicht mehr nur Konsumenten sind, wie es die traditionelle mikroökonomische Theorie sieht, sondern vielmehr auch "Co-Produzenten" von Sachgütern und Dienstleistungen (Kortzfleisch 1997). Aufgrund der Nutzung neuer Informations- und Kommunikationssysteme durch die Bürger sind folgende Veränderungen zu erwarten:

1. *Verbesserte Transparenz der Vorgänge:* Direkter, individueller Zugang zu Informationen erhöht die Transparenz der Vorgänge und Vorkommnisse, die alle wirtschaftlichen, sozialen, politischen, öffentlichen und kulturellen Angelegenheiten betreffen.
2. *Verbesserung der Kommunikation:* Die Nutzung der Mehrwertdienste führt zu einem generellen Wandel der Kommunikationsstruktur, insbesondere zu einem quantitativen und qualitativen Wandel des Kommunikationsverhaltens der Bürger.
3. *Veränderung des Transaktionsverhaltens:* Das Transaktionsverhalten ändert sich durch die Mehrwerte von elektronischen Marktplätzen.
4. *Erhöhung der Mobilität:* Die Nutzung der elektronischen Marktplätze fördert die Mobilität des Faktors Arbeit, woraus sich Verschiebungen in der regionalen Beschäftigungsstruktur ergeben.
5. *Starke Partizipation am öffentlichen Leben:* Die Nutzung elektronischer Marktplätze führt zu einer stärkeren Partizipation der Bürger am öffentlichen Leben.

3 Erhebungsverfahren

Die Ausgestaltung der Fragebögen wurden nach den Vorgaben von Diekmann, Dillmann und Laatz durchgeführt (Diekmann 1995, Dillmann 1978, Laatz 1993). Als Erhebungsverfahren wurde die schriftliche Befragung gewählt. Im Bereich Wirtschaft und Verwaltung konnte eine Vollerhebung, bei den Bürgern eine repräsentative Befragung durchgeführt werden.

4 Ergebnisse

Die ersten Ergebnisse haben gezeigt, daß bereits heute 27 % der Unternehmen das Internet nutzen oder dessen Einsatz planen (siehe Abb. 2). Zur Zeit beschränkt sich jedoch der Einsatz auf die einfache Präsentation im Web. Verhalten werden die Möglichkeiten eines weitergehenden Engagements auf

elektronischen Märkten, wie z. B. elektronischer Einkauf oder die Integration des Internet in das Bestellwesen, gesehen. Die Studie hat aufgezeigt, daß sich die unternehmerische Nutzung elektronischer Mehrwertdienste noch in einem Anfangsstadium befindet.

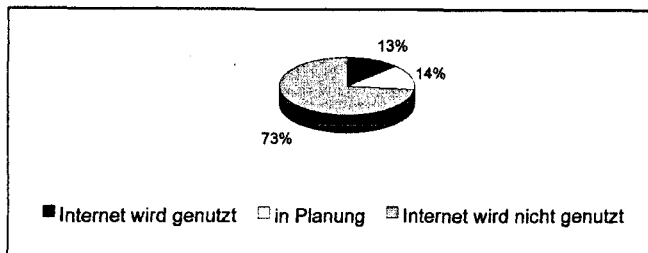


Abb. 2: Internetnutzung der Pfaffendorfer Unternehmen

Sowohl die Verwaltungsmitarbeiter (75 %) als auch die Bürger stehen den zukunftssträchtigen elektronischen Verwaltungsdiensten durchaus positiv gegenüber. Die Einführung eines elektronischen Bürgerinformationssystems soll nach Meinung der Verwaltungsmitarbeiter Orientierungsinformation wie z. B. Öffnungszeiten, Zuständigkeiten oder mitzubringende Unterlagen in den Vordergrund stellen. Bereits 60 % der befragten Bürger können sich vorstellen, Informationen der Stadtverwaltung elektronisch abzurufen und dies, obwohl zur Zeit 75 % der Bürger sich mit Hilfe des Telefons ihre Verwaltungsinformationen beschaffen (siehe Abb. 3).

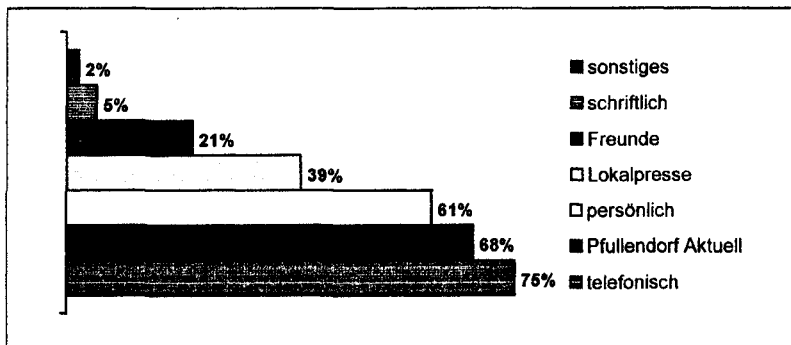


Abb. 3: Informationsquellen für Verwaltungsangelegenheiten

Dabei kommt es den Bürgern nicht nur auf Orientierungsinformationen, sondern vielmehr auf Informationen über aktuelle politische Themen, Beschlüsse und Projekte der Stadt an. 52 % der befragten Bürger wollen in Zukunft elektronische Formulare von ihrem PC oder von öffentlichen Terminals aus

nutzen. 41 % der Unternehmen stehen dem Online-Verwaltungsgeschäft ebenfalls aufgeschlossen gegenüber.

Die Auswertung der Bürgerbefragung hat gezeigt, daß das Informationsbedürfnis der Bürger im wesentlichen von den bestehenden Informationsangeboten gedeckt wird, wenngleich bei einigen Bereichen, wie z. B. beim System der öffentlichen Versorgung unbefriedigende Informationslücken bestehen. Abrufinformation sollte um eine Kombination aus elektronischen Informations- und Kommunikationsangeboten wie z.B. Email angereichert werden. Die meisten Bürger können sich noch nicht vorstellen, auf einem elektronischen Marktplatz einzukaufen, zu bestellen, zu buchen, zu reservieren oder zu bezahlen. Die Bereitschaft, solche Transaktionen auf elektronischem Weg durchzuführen, ist noch sehr gering. Es hat sich gezeigt, daß den Bürgern noch das notwendige Vertrauen in diese Online-Dienste fehlt (siehe Abb. 4).

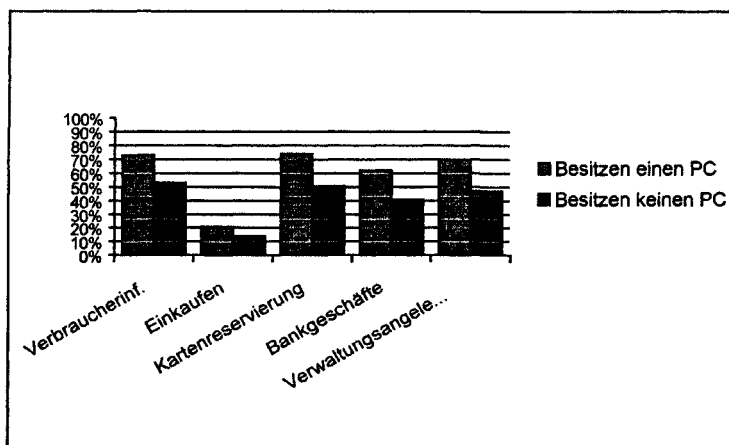


Abb. 4: Erledigung über Online-Dienste

5 Ausblick

Zwei Jahre nach der Erstbefragung wird die zweite Erhebung im Herbst 1998 durchgeführt. Während dieser Zeit wurde ein elektronischer Marktplatz namens „KISS“ stellvertretend für „Kommunikation, Information, Shopping und Service“ in Pfullendorf aufgebaut. An verschiedenen öffentlichen Plätzen wurden multimediale Computerterminals aufgestellt, die einen kostenlosen Zugang zu lokalen und regionalen Informationen garantieren. Durch diese Maßnahme sollen die Bürger zu einer vermehrten Nutzung dieses elektronischen Marktplatzes angeregt werden. Im nächsten Schritt werden transaktionsorientierte Mehrwertdienste wie beispielsweise das Bezahlen im Internet mit Chipkarte oder mit Kreditkarte aufgebaut. Für die Unternehmen, die ihre Geschäftsbeziehungen in Zukunft virtuell abwickeln wollen, ist es

notwendig, daß durch den Einsatz elektronischer Signaturen rechtsverbindliche Transaktionen ermöglicht werden. Die Implementierung des Open Trading Protokolls (OTP) für sichere Transaktionen und Geschäftsprozesse im Internet, von der Bestellung bis zur Lieferung, ist ein weiterer Schritt in der Entwicklung dieses elektronischen Marktplatzes. Es hat sich gezeigt, daß eine grundsätzliche Bereitschaft der Bürger, Unternehmen und Verwaltung zum Einsatz und zur Nutzung neuer Multimedia- und Kommunikationssysteme vorhanden ist. Jedoch ist die Etablierung von Vertrauen in elektronische Prozesse, gestützt durch moderne Informationssysteme, eine Erfolgsnotwendigkeit. Die Ergebnisse der zweiten Untersuchung werden zeigen, ob und wie die durchgeführten Maßnahmen eine Veränderung im Untersuchungsgebiet erbracht haben.

6 Literatur

[Bell 1976]

Bell, Daniel: *Die nachindustrielle Gesellschaft*. Frankfurt, New York: Campus, 1976

[Bievert 1994]

Bievert, B.; Monse, K.; Gatzke, M.; Reimers, K.: *Digitaler Dienst am Kunden - Informationstechniken an der "Kundenschnittstelle"*. Berlin, 1994

[Bullinger 1995]

Bullinger, Hans-Jörg; Brettreich-Teichmann, Werner; Fröschle, Hans-Peter: *Das virtuelle Unternehmen: Koordination zwischen Markt und Hierarchie*. In: Office Management 12, 1995

[Diekmann 1995]

Diekmann, Andreas: *Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen*. Reinbek: Rowohlt, 1995

[Dillmann 1978]

Dillmann, D. A.: *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley, 1978

[EM 1991]

Elektronische Märkte. St. Gallen: Hochschule / Institut für Wirtschaftsinformatik, 1991. Nr. 1, Sept. 1991

[Grabow 1996]

Grabow, Busso; Korte, Werner B.: *Telematik, Teledienstleistungen und Kommunalpolitik*. Köln: Deutsches Institut für Urbanistik (Difu, Aktuelle Informationen), 1996

[Hensel 1990]

Hensel, Matthias: *Die Informationsgesellschaft, Neuere Ansätze zur Analyse eines Schlagwortes*. München: Verlag Reinhard Fischer, 1990

[Horvath 1996]

Horvath, Peter; Gleich, Ronald; Voggenreiter, Dietmar: *Controlling umsetzen*. Fallstudien, Lösungen u. Basiswissen. Stuttgart: Schaeffer-Poeschel, 1996

[Klein 1994]

- Klein, Stefan: *Entwicklungstendenzen elektronischer Märkte: von High-Volume Commodity Märkten zu Konsumentenmärkten*. St. Gallen: Hochschule St. Gallen, 1994
- [Kortzfleisch 1997]
Kortzfleisch, Harald F.O. v.; Nünninghoff, Klaus; Winand, Udo: *Ansatzpunkte für die Entwicklung haushaltsgerechter Benutzungsoberflächen beim Einsatz neuer Medien- und Kommunikationssysteme an der Kundenschnittstelle*. In: *Wirtschaftsinformatik*, Heft 3. Wiesbaden: Vieweg, 1997
- [Kubicek 1995]
Kubicek, Herbert: *Zur informierten Stadt durch elektronische Informationssysteme? Zwischenbericht*. Bremen: Universität Bremen/Forschungsgruppe Telekommunikation, 1995
- [Kuhlen 1995]
Kuhlen, Rainer: *Informationsmarkt - Chancen und Risiken der Kommerzialisierung von Wissen*. Konstanz: UVK, 1995
- [Kuhlen 1996]
Kuhlen, Rainer: *Zur Virtualisierung von Regionen durch elektronische Marktplätze*. *Berichte der Informationswissenschaft* 79-96. Konstanz: Universität Konstanz, 1996
- [Kuhlen 1996b]
Kuhlen, Rainer: *Elektronische Märkte in der Informationsgesellschaft oder: die Informationsgesellschaft als elektronischer Markt*. In: Glowalla, U.; Schoop, E. (Hg.), *Deutscher Multimedia Kongress 1996*. Berlin u. a.: Springer, 1996, S. 43-48
- [Komon 1993]
Kumon, Shumpei; Aizu, Isumi: *Co-Emulation: The Case for a Global Hypertext Society*. In: Harasim, Linda M.: *Global Networks, Computers and International Communication*. Cambridge, London: The MIT Press, 1993
- [Laatz 1993]
Laatz, Wilfried; Deutsch, Harri: *Empirische Methoden: Ein Lehrbuch für Sozialwissenschaftler*. Thun, Frankfurt: 1993
- [Palass 1997]
Palass, B.; Preissner, A.; Rieher, J.: *Managen in Zeiten des Cyberspace*. In: *Manager Magazin* Nr. 3, 1997
- [Rieser 1992]
Rieser, Ignaz: *Informationstechnologien und Organisation*. In: *Die Unternehmung* 5, 1992
- [Tapscott 1996]
Tapscott, Don: *Die digitale Revolution: Verheißungen einer vernetzten Welt - die Folgen für Wirtschaft, Management und Gesellschaft*. Wiesbaden: Gabler, 1996

Interface for Navigation and Database - Querying via WWW

A.L. Heuer, K. Siemonsen, T. Engel, C. Meinel

Institut für Telematik
Bahnhofstraße 30-32, 54292 Trier
{Heuer, Siemonsen, Engel, Meinel}@ti.fhg.de

Contents

Summary

1. Introduction

2. The Design of the System

2.1 The Metadatabase

2.1.1 The Meta Information

2.1.2 The Database Information

2.1.3 Atomic Storage of Query-Components

2.1.4 The Navigation Structure

2.2 The Server

2.3 The Administration Tools

3. Possible Employment of the System

4. Conclusion and Outlook

Literature

Summary

We describe the concept and the implementation of a Navigation and Database Interface (NDI) on the World Wide Web. The prototype combines the storage of database queries and navigational structures, both made more valuable by comprehensive meta information. The multilingual design of the meta information set prepares the system for use in an international environment.

1. Introduction

As described in (Benn 1998) there is an enormous number of systems, scientific or commercial, trying to make databases available on the WWW. The most common way is using forms and CGI-scripts. Although this is a very easy solution, it becomes more and more unmanageable if the number of queries

has to be extended or modified frequently. The number of scripts will increase with the number of queries. The performance is rather poor and problems occur as the number of accesses rises.

Another common approach uses macro-HTML. Documents include special tags that are preprocessed by the server before sending the document to the client, so that the tables are generated on the fly. The results are put into predefined HTML-pages either by CGI scripts or by an added server functionality. This method is implemented easily and allows better maintenance than using CGI-scripts that mix executable code with database queries and layout. Some database vendors implement HTTP interfaces in their database engines. HTTP requests to such databases are processed directly. However, this method will only work on a proprietary basis. Each database vendor will create his own interface.

Finally there is the more complex approach of using distributed services. Using IIOP or other protocols, applications - distributed by application servers to the clients - connect to objects implementing business logic in order to manipulate data. The client only displays data, while database interaction uses appropriate services provided by database applications.

Our approach is similar to a database engine implementing HTTP. We built a server that delivers navigational information enriched with meta information. Such meta information are, as described below, a speaking name, a comment, a number of keywords and an additional HTML page containing detailed information. Meta information is stored in different languages. Therefore the same navigational structure can be used in multilingual environments. On the client side users receive dynamically generated HTML pages containing links. Each link leads either to another navigational page lower in the hierarchy or represents a database query on a relational database. Meta information is provided in either case. If users choose a database query link, the according query will be executed. The result set will be returned as a table.

2. The Design of the System

The system architecture is quite simple. A relational (Codd 1990) database, the so called metadatabase, is used to store information about other relational databases. An interface based on this metadatabase provides access to the databases known to the system.

The complete system, as shown in Fig. 1, actually consists of four components on the server side. Firstly there is the metadatabase itself. For example, it stores information about relational databases (DB1, DB2, ...). Furthermore there is the HTTP-server-like program (NDI) that provides the interface. It displays the navigational structures and executes the select-statements. Finally there are two tools for administration. One (Admin 1) is being used to maintain the information about the managed databases. The other one (Admin 2) allows to generate queries and navigation paths. Platform independence is aspired by doing all coding with Java. Therefore the server component and the administration tools are expected to run on most systems. Database connections are managed by JDBC drivers (Reese 1997). Because of our decision to use simple HTML pages submitted to the client via HTTP, every common WWW-Browser can be used as a front-end to the system.

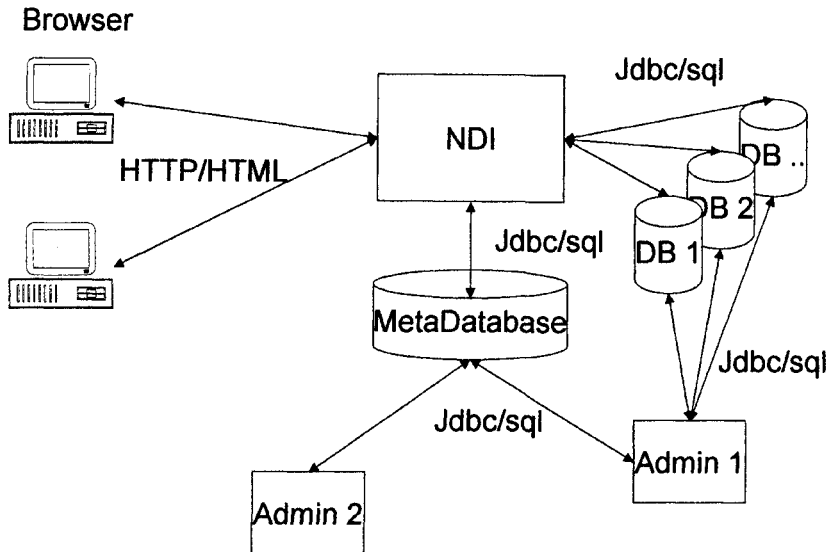


Fig. 1 The system architecture of the Navigation and Database Interface (NDI). The Browser as a client, and the NDI as Server that connects to databases using information from the metadatabase. Furthermore the two Admin tools to manage the alignment between the metadatabase and the maintained databases (Admin 1) as well as the navigational structure (Admin 2)

2.1 The Metadatabase

The data needed to provide a navigational interface to the client is stored in a relational database. This database contains different kinds of information. Firstly there is the information about the databases known to the system. Secondly there are navigational structures. Furthermore the metadatabase retains the queries to be executed on the known databases. Finally it stores metadata to enable a user-friendly informational interface to the databases with their contents.

The metadatabase consists of over thirty tables managing the mentioned information. The following paragraphs describes the different kinds of information in detail.

2.1.1 The Meta Information

In our system meta information (Fig. 2) is an additional description for all objects stored in the metadatabase. This means there is description maintainable for single columns as well as for complete queries. Such meta information is at first a speaking name that is presented to the user for each object. For example, the system allows administrators to give each column a name describing the content of the column to the common user. This name can be defined simultaneously in any language maintained by the system. In addition to the speaking names it is possible to enter multilingual comments. They allow a

short description of the objects they are linked to. For example, they may be displayed in the status line of the Browser with JavaScript, if the user moves the mouse pointer over the corresponding link. Finally the administrators may use HTML pages served by a common webserver for detailed description.

Meta Information

Language 1	Language 2	Language n
Speaking Name	Sprechender Name	...
Comment	Kommentar	...
Http://server/help.html	Http://server/hilfe.html	...

Fig. 2 The metadata managed for each object stored in the metadatabase.

The system stores only the URLs of these detailed help-pages. Therefore common techniques can be used to create and store the pages. Of course such detailed pages can also be differentiated by language. Furthermore such description pages may contain links to other relevant database contents. This adds an additional navigation layer, that is not maintained by the NDI anymore. In order to ease finding of an object a set of multilingual keywords can be attached to the objects.

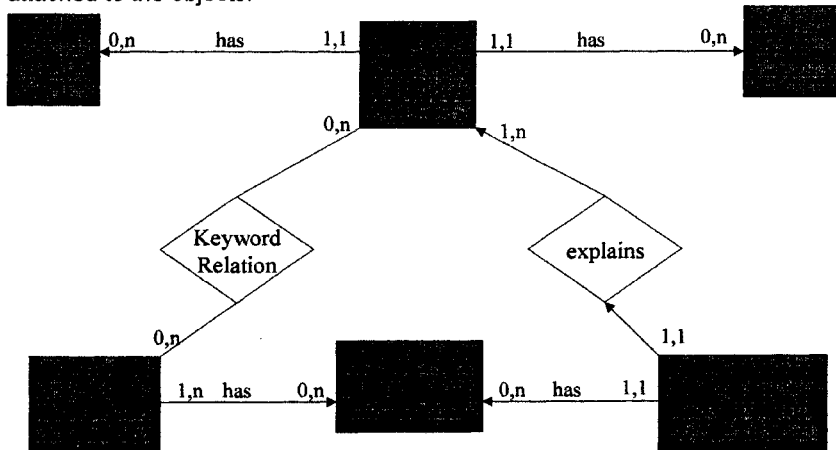


Fig. 3 The meta object with its relations. Other objects that need meta information inherit from this object.

In the data model a meta object (Fig. 3) was designed that contains the set of meta information. Each other object that needs access to metadata inherits from this object .

2.1.2 The Database Information

In order to access tables or views of any other relational database with the NDI, the metadatabase has to store information about their structures. This means that each table or view with its columns and its data formats has to be referenced in the metadatabase. Parameters to access those databases are stored as well as their type, the corresponding hostname and the port of the database server.

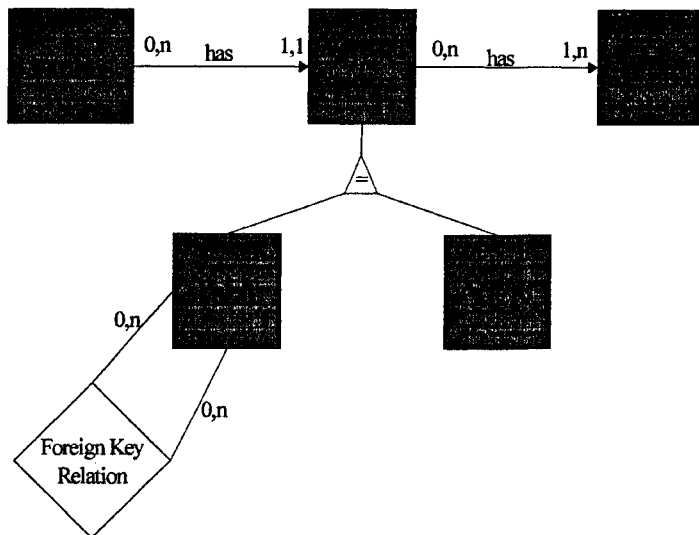


Fig. 4 The relationship of the entities relevant for the access of the maintained databases. Because of their similarities, tables and views are described in the model as virtual tables. Tables and views only extend those virtual table objects.

While this approach makes it difficult to handle constantly modified database structures, it is satisfying for databases that are set up once and modified only very seldom. Of course a modification in the sense of extension is easier to handle than removing columns or tables.

2.1.3 Atomic Storage of Query-Components

A common SQL (Date 1997) statement consists of several components. The NDI system does not store queries as strings. Each component is stored by reference separately. This means a select statement is broken down on column and table / view basis. Because the metadatabase does not use strings to store queries, it is very easy to modify a query afterwards. Furthermore, if some table

is modified or a column is removed, there is no need to parse query strings. One can directly search in the sets of query components. Since each query consists of a set of components, the system creates query strings dynamically. If one wants to use the same query on a different database system, the syntax can differ. The dynamic creation of query strings allows to adapt the syntax. Furthermore the atomic storage of the query components can be used for query inheritance in case of a further drill down. For each inheritance level the administrator can define which component he wants to add to or remove from the actual basic query. So such a basic query can be extended until the resultset is the desired one.

Restriction of queries can be done by three different kinds of values. Firstly there are static values. They are set at the time of query creation. Furthermore the system provides two kinds of dynamic values. Date and time periods, e.g. today, yesterday or last 20 days are generated for restriction dynamically before the query is sent to the database. Finally there are text fields that have to be filled out by users at runtime and lists, from which users may select entries. In such a case of interaction with users a form has to be shown before the execution of the query.

2.1.4 The Navigation Structure

Besides the database dependent information the system has to store navigational information (Fig. 5). Therefore a hierarchical structure was chosen. The system stores nodes and edges.

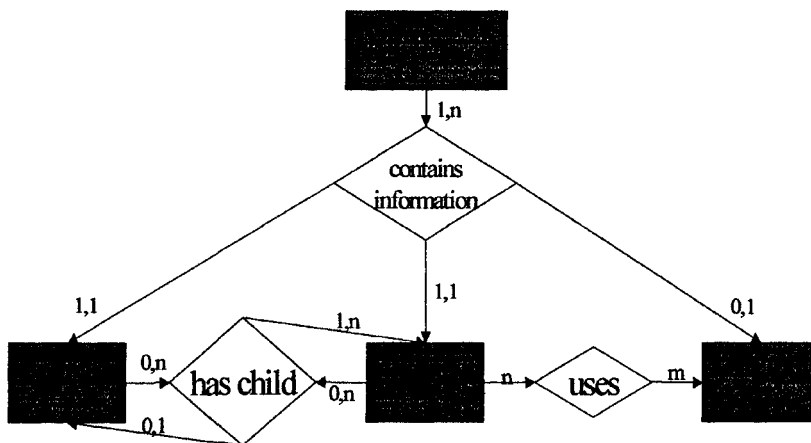


Fig. 5 The relationship between nodes and statements. Each node can have one parent node. A node may have child nodes and statements. Each statement can be linked to an arbitrary number of nodes. A statement is executed by the database it was built for. For nodes as well as statements and the original databases meta information is available.

For ease of maintenance each node may have several child nodes but only one parent node. A node contains a set of meta information, as described above.

Connecting the nodes with edges creates a hierarchy which allows an upward and downward navigation in the structure. Equivalent to the directory structure in file systems, nodes may have child nodes that would represent subdirectories. To each node a number of queries (queries could be associated with the files in file systems) can be attached. Those queries may aim at different databases. So in the navigational structure it is possible to group queries by the content they deliver, independent of the database that stores the data. Users will not need to know which database stores the data they require. They will follow the content dependent paths.

One query object can be linked to many nodes, eventually in different contexts. So it is possible to build more than one path to the same set of queries. Therefore different views on the data can be generated. This allows to adapt the system for different groups of users. The system does not build a content dependent navigational structure automatically. Administrators have to develop such structures manually. In the node and query space the search by keywords that can be attached to them would be a kind of free navigation.

2.2 The Server

Using Java the NDI server was implemented platform-independent. The server program (Sridharan 1997) listens for HTTP requests on a specified port. It parses the requests and extracts the required information from the metadatabase. Either the system generates then a new navigation page that is presented to the user, or it executes a query on a database. In this case the received resultset will be shown to the user in a HTML-Table.

In the navigation pages each node object is represented by its speaking name. The name is connected via a hyperlink to the child nodes and the attached queries.

The server program (Fig. 1, NDI) runs multithreaded and is capable to connect to several different database engines (from different vendors) at once. The connection to all databases is done by JDBC. Therefore the server can only connect to databases which provide JDBC drivers that make them accessible via internet. Since nearly each bigger vendor provides a driver to his databases, this is no real restriction. In order to adapt to user requests, the server logs each database query. This may help to specify existing and further needed queries.

2.3 The Administration Tools

As mentioned above (Fig. 1) there are two tools that were designed in connection with the Navigation and Database Interface. Both tools are quite platform independent, since their coding was done with Java. They connect to the databases with JDBC in the same manner as the server.

The first one (Fig. 1, Admin 1) is needed to import the structures of the databases to be maintained by the system. It reads the system tables of the databases and extracts the required information from them. Since the system tables differ for each database type, this process requires some adaptational work, if a new database is to be imported. When JDBC drivers provide a complete set of database metadata the import process can be standardized. As described above each table, view and column of the maintained databases is referenced in the metadatabase. Furthermore meta information may be added during the import process. Besides this import functionality, certain consistency

checks have to be implemented. The complete system will only work, if the information about the original databases concerning tables and columns is correct.

After an import the available information can be managed with a second tool (Fig. 1, Admin 2). This tool allows to create the hierarchical structures for navigation described above.

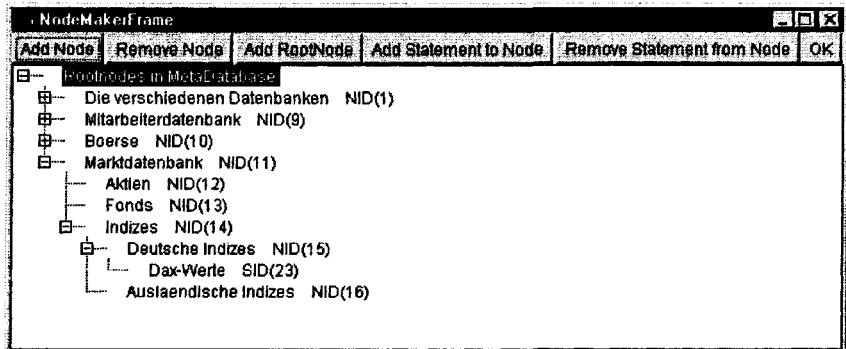


Fig. 6 The window with the tree view that displays the navigational structure of the NDI. The window can also be used for editing.

The navigational structure is shown in a tree view (Fig. 6) that may be edited with the functions the window provides. A meta information window (Fig. 7) gathers the information associated with a node or query. The information can be entered in several languages at the same time.

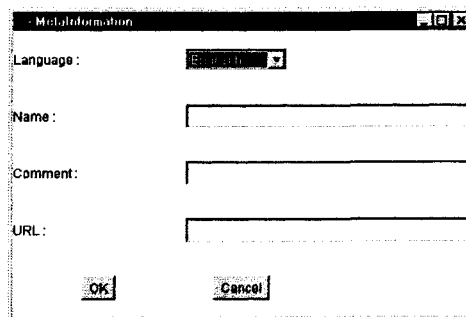


Fig. 7 The window that gathers the meta information for a node or statement. The choice in the top switches the language.

Simple queries can be designed with a query design window (Fig. 8).

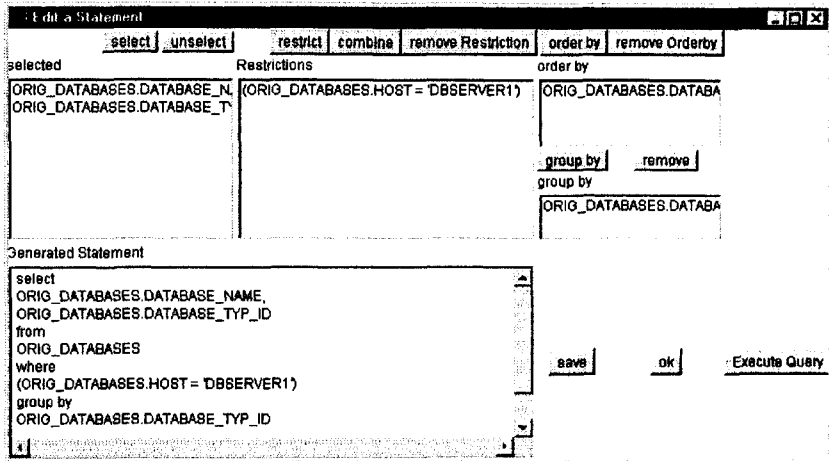


Fig. 8 The window used to generate simple select statements.

Each query may be attached to several nodes. This enables different navigational structures leading to the same data. So users with different preferences concerning navigation can be satisfied. While the first tool is implemented as an application, the second one will run as an Applet as well. This allows database administrators to create their queries for the databases maintained by themselves via Internet.

3. Possible Employment of the System

The main aim of the NDI is to increase the availability of existing databases. The system provides one general user interface for databases of different vendors. It does not bother users with the need for knowledge of each databases existence, content and handling. Users just select what kind of information they need by following the predefined paths built by experienced people (administrators). The system will lead them to the available database queries. So Users do not even need to know anymore, which database answers their queries. They just receive a table containing a result set. If they need to gather further information, they may read the comment or the additional help page belonging to the query they executed. Because of its multilingual implementation and the meta information it provides, the system could be employed for the documentation of databases, too.

The most important use of the NDI is probably the intranet of a multinational company or an international organization. Even public databases, e.g. in environmental protection (Siemonsen 1998) could possibly use the system. Usually such companies and institutions own several "older" databases with very valuable data. Often it is unreasonable to migrate such systems to more up to date database engines, since there are still some applications running, that require the existing database engines. Up to now, very often one has to use special, database-dependent, difficult to use proprietary tools, which are very often platform-dependent, in order to extract information from these databases.

4. Conclusion and Outlook

In this paper we described the concept and prototype implementation of a Java based and therefore platform- independent database interface on the World Wide Web. This interface combines navigation and database queries. Both are enriched with multilingual metadata. Of course the system will not be satisfying for people being used to access databases directly with SQL in a proprietary environment. But it works very well for a great percentage of common users that have to execute nearly the same set of queries on several databases every day. Furthermore it supports users with their search for information, of which they do not know where it is stored.

Database administrators can generate statements satisfying users needs and provide them by using the interface.

While still being a prototype, the system can be improved in several ways. In order to create a more general way to run the server component, we intend to implement the server service described above as a servlet. Then each webserver capable of running Java servlets could be extended with the Navigation and Database Interface (NDI). At the moment only database querying is provided. If we decide to manage sessions and implement user registration, it is also feasible to allow data manipulation. In the near future we plan to extend the system in regard of URL managing. Then not only database queries can be attached to nodes, but arbitrary URLs pointing to content related data sources. This will allow a navigation in an even more open information space. In contrast to a search engine the system would deliver URLs that have been reviewed and sorted by experienced people.

Literature

[Benn 1998]

Benn, W. Gringer, I. : *Zugriff auf Datenbanken über das World Wide Web*, Informatik Spektrum 21:1-8(1998), Springer Verlag 1998

[Codd 1990]

Codd, E.F. : *The Relational Model for Database Management: Version 2*. Reading, MA: Addison-Wessley 1990

[Date 1997]

Date, C.J., Darwen,H.: *A Guide to SQL Standard*. 4. Auflage. Reading, MA: Addison-Wessley 1997

[Reese 1997]

Reese, G.: *Database Programming with JDBC and Java*, O'Reilly 1997

[Siemonsen 1998]

Siemonsen, K. Heuer, A. Engel, T. and Meinel C.
Ein Web-basiertes Navigationssystem für allgemeine, heterogene Datenbanksysteme, Proceedings of 1. Workshop Hypermedia im Umweltschutz, Ulm (1998) ,p. 83-86

[Sridharan 1997]

Sridharan, P. : *Advanced Java networking*, Prentice Hall 1997, p.67-75

Elektronische Kommunikationsforen – das Beispiel UNESCO VF-INFOethics

Rainer Kuhlen, Anja Odenthal

Informationswissenschaft, Universität Konstanz
Fach D87, D-78457 Konstanz
{kuhlen/odenthal}@inf-wiss.uni-konstanz.de

Inhalt

- 1 Faktor Kommunikation
- 2 Elektronische Kommunikationsforen
- 3 Design von ComFor für VF-INFOethics
- 4 Ergebnisse
 - 4.1 VF-INFOethics - einige Ergebnisse
 - 4.2 ComFor – einige Ergebnisse
 - 4.3 Geplante Weiterentwicklung

Zusammenfassung

Momentan etablieren sich auf vielen WWW-Sites elektronische Kommunikationsforen - die zwar nicht die klassischen Formen der Kommunikation ersetzen, sie aber in einigen Punkten ergänzen. Bei der Informationswissenschaft wurden in den letzten Jahren einige Diskussionsforen entwickelt und eingesetzt. Beispielhaft wird das System ComFor vorgestellt, mit dem das Forum VF-INFOethics betrieben wurde. Wichtigstes inhaltliches Ergebnis war dabei, daß Kommunikationsforen gewissermaßen Instrumente kommunikativer Bildung darstellen; die Teilnehmer lernen schnell, daß es je nach kulturellem und persönlichem Hintergrund individuelle Kommunikationsstile gibt, auf die sehr sensibel reagiert werden muß – Kommunikationsforen können also eine wechselseitige Annäherung und Respektierung der Teilnehmer untereinander erzielen. In technischer/konzeptioneller Hinsicht ist die Orientierung und Navigation wichtigster Gesichtspunkt bei der Entwicklung von Kommunikationsforen.

1. Faktor Kommunikation

Kommunikation rückt unter den Bedingungen globaler Netzwerke und elektronischer Marktplätze erneut ins Zentrum informationswissenschaftlichen

Interesses und ergänzt die Informations-, Präsentations- und Transaktions-sichten um den wichtigen Aspekt, Konsens über wichtige Fragen im Austausch mit den Positionen anderer zu erzielen. Daß Kommunikation zentrales Element wissenschaftlicher Tätigkeit überhaupt ist, ist seit langem bekannt und als Erfolgsfaktor unumstritten. Der Erfolg amerikanischer Wissenschaft wird zu großen Teilen auf die traditionell große Kommunikationsbereitschaft erklärt, die seit über 30 Jahren auch durch die technische Vorreiterrolle in Sachen technischer Kommunikation (von ARPA zum Internet und zur NII) gefördert wird und Wettbewerbsvorteile kontinuierlich sichert. Aber auch in allen anderen Bereichen wird Kommunikation als wesentliches Element bei der Ausgestaltung von Aktivitäten auf elektronischen Marktplätzen gesehen. War die erste Phase des Internetauftritts von kommerziellen und politisch-administrativen Organisationen weitgehend noch durch das Ziel der Präsentation von Leistungen bestimmt, die zweite Phase durch Interaktionskomponenten wie Auswahl und Bestellung, die dritte durch Transaktionskomponenten wie elektronische Zahlungsformen, so rückt gegenwärtig die kommunikative Funktion in den Vordergrund – die anderen Funktionen nicht ablösend, aber diese doch ergänzend bzw. sie auf eine neue Stufe hebend. Dabei sind die Interessen durchaus divergierend: bei wissenschaftlicher Kommunikation steht – alles sicherlich idealtypisierend – der Austausch von Information und das Gewinnen von Anerkennung im Vordergrund, im Wirtschaftsbereich der Aspekt der Kundenbindung und der Möglichkeit eines intensivierten 1-to-1-Marketing, in der Verwaltung der Wunsch nach Bürgernähe und funktionaler Effektivität und auf den allgemeinen Publikumsmärkten die Attraktivität des bequemen Kontaktes mit Freunden, Verwandten, aber auch ganz neuen, bislang unbekannten Partnern sowie das allgemeine Bedürfnis nach Unterhaltung.

2. Elektronische Kommunikationsforen

Entsprechend der erkannten Bedeutung von Kommunikation und der universal sich durchsetzenden Verwendung der Web-Technologie entstehen auf vielen Web-Sites elektronische Kommunikationsforen als selbständige und als in andere Angebote integrierte Anwendungen¹. Foren etablieren sich also als selbständige Dienstleistung neben den bisherigen Kommunikationsdiensten wie Email, Bulletin Boards/ListServer, Chats, MUDs oder auch Tele-/Video-

¹ Einige Beispiele:

- a) **Alleinstehende Foren.** Z.B. "Ethics Updates " (<http://ethics.acusd.edu/scripts/WebX.dll>), Diskussionsforum "Wahl 1998" (<http://www.rudow-net.de/wahl98/index.html>)
- b) **E-Journals.** Z.B. "Science's Next Wave" (<http://www.nextwave.org/>) "Review of Information Science" (<http://www.inf-wiss.uni-konstanz.de/RIS/>);
- c) **Virtuelle Konferenzen.** Z.B. WSC3 (3rd On-line World Conference on Soft Computing in Engineering Design and Manufacturing) (<http://www.cranfield.ac.uk/wsc3/>)
- d) **Distance Learning – Anwendungen.** Z.B. Lotus LearningSpace (<http://www.lotus.com/products/learningspace.nsf>), College of the Sequoias Distance Education Program (<http://zeus.sequoias.cc.ca.us/distance/index.html>)

Conferencing. Foren können wie Chats oder MUDs auch synchron organisiert sein bzw. können synchrone Teile enthalten, zeichnen sich aber in der Regel durch asynchrone Kommunikation aus, d.h. erlauben zwar auch eine unmittelbare, fast real-time Reaktion, gestatten aber – was im globalen Maßstab unabdingbar und zur Erhöhung der Qualität der Beiträge wichtig ist – auch eine verzögerte, informationell abgesicherte (und nachgedachte) Reaktion. Foren stellen eine wesentliche Ausprägungsform partizipativer Massenmedien² dar.

Elektronische Kommunikation – darin den meisten medialen Leistungen vergleichbar – ersetzt nicht die klassischen Formen der Face-to-face-Kommunikation oder andere traditionelle Formen des Austauschs (z.B. Publikationen, Konferenzen), substituiert sie aber partiell und erlaubt neue Formen des Austauschs, die bislang nicht möglich waren. Wir machen die folgenden Mehrwerte von elektronischen Foren aus:

- Elektronische Kommunikationsformen gestatten den Austausch von Information zwischen Leuten, die im „realen“ Leben kaum eine Chance haben, sich zu treffen. Elektronische Kommunikationsformen erleichtern insbesondere den Kontakt zu Angehörigen anderer Kulturen und können somit einen wichtigen Beitrag zur Bildung von Toleranz und multikultureller Bewußtheit leisten.³
- Elektronische Kommunikationsformen, die nach Hypertextprinzipien organisiert sind, gestatten eine flexible Navigation in heterogenen Informationsmaterialien und können sich daher in eine allgemeine, auf das jeweilige Themengebiet bezogene Wissensplattform entwickeln.
- Elektronische Kommunikationsformen können Vorteile synchroner, aber vor allem asynchroner Kommunikation nutzen, gestatten also quasi unmittelbare Reaktionen, aber auch reflektierte verzögerte.
- Elektronische Kommunikationsformen bieten in vielen Bereichen eine vergleichsweise kostengünstige Substitution herkömmlicher Kommunikationsformen⁴, vorausgesetzt die technische Infrastruktur ist schon einmal vorhanden.
- Elektronische Kommunikationsformen erlauben eine flexible Anpassung an unterschiedliche Benutzererwartungen und –sichten und können flexibel

² Definition partizipatives Massenkommunikationssystem: "Wenn die Anzahl der potentiellen Sender und Empfänger nicht begrenzt ist und entsprechend die über das Kommunikationssystem verfügbaren Kommunikationsangebote potentiell jeden zugänglich (d.h. öffentlich) sind, [...] dann könnte man von *partizipativen oder interaktiven Massenkommunikationssystemen* sprechen. Diese zeichnen sich dadurch aus, daß Kommunikation nicht einseitig von einem Sender zu einem Empfänger, also nicht "one-to-many", sondern auch wieder zurück, d.h. "many-to-many", verläuft und damit die Rollentrennung von Kommunikator und Rezipient auflöst." [Höflich 96]

³ (Fowler 95)

⁴ z.B. im Bereich virtueller Konferenzen oder im Bereich Wissenschaftspublikationen (E-Journals)

nach unterschiedlichen Benutzergruppen mit verschiedenen Rechten des Ansehens/Lesens und Schreibens organisiert werden.

- Elektronische Kommunikationsformen bieten eine Form der Konsensbildung zwischen Gruppen, die sich nach systematischen oder topologischen (regional/global) Kriterien definieren.

In der Konstanzer Informationswissenschaft sind in den letzten Jahren eine Vielzahl elektronischer Kommunikationsforen eingerichtet worden. Diese Foren basieren auf zwei unterschiedlichen Softwaresystemen: zum einem dem KHS (Konstanzer Hypertext System) unter Unix und zum anderen ComFor unter Windows NT. KHS-Foren wurden eingerichtet unter anderem für die interne Mitarbeiterkommunikation, für eine elektronische Zeitung in der Electronic Mall Bodensee, für die elektronische Zeitschrift RIS, für die Kommunikation innerhalb von ASIS-Europe oder zwischen der Konstanzer Informationswissenschaft und dem Karlsruher Fachinformationszentrum. ComFor-Foren wurden unter anderem für die interne Kommunikation im Projektmodul "Medienassistenten" des BMBF-Projektes DeMeS (Development of Media Services), für mehrere Kurse im Wintersemester 97/98 und im Sommersemester 98 sowie bei Projektpartnern⁵ eingerichtet.

Wir wollen hier die Möglichkeiten der elektronischen Kommunikationsforen am Beispiel von VF-INFOethics demonstrieren, eines elektronischen Kommunikationsforums, das für die UNESCO eingerichtet wurde, um Probleme der Informationsethik in einem globalen Maßstab unter ausgewählten Experten zu diskutieren und um daraus Empfehlungen für die UNESCO abzuleiten.

Bei dieser Anwendung interessiert uns vor allem das angeführte Kriterium der Konsensbildung, gerade auch in Situationen der interkulturellen Kommunikation, wie sie für die Internetdienste und elektronische Marktplätze und insbesondere für Kommunikationssituationen der UNESCO die Regel ist. Kommunikationsforen bauen hierarchische Strukturen dadurch ab, daß auf Kommunikationsforen in erster Linie die eingebrachten Beiträge zählen, nicht der Status, das Vor-Prestige⁶.

Kommunikationsforen sind nach unserer Einschätzung und Erfahrung aber auch ein wichtiges Element der wechselseitigen Annäherung und der Respektierung der „Anderen“. In einem global organisierten Forum wie VF-INFOethics wird sehr schnell gelernt, wenn man denn nicht gänzlich unsensibel ist, daß es je nach kulturellem und individuellem Hintergrund individuelle Kommunikationsstile gibt, auf die sehr sensibel reagiert werden muß. Interkulturelle Unterschiede schließen, wie es sich zeigt, keineswegs Verständigungsmöglichkeiten aus. Die Kommunikation glückt aber nur dann, wenn die eigenen Standards nicht absolut gesetzt werden und wenn die Bereitschaft der Anerkennung des „Anderen“ und der Bereitschaft, Konsens durch Annäherung, nicht durch Mehrheiten zu erzielen, gegeben sind.

⁵ Insbesondere bei der Informationsmanagement GmbH, Stuttgart und dem Steinbeis-Transferzentrum IMAC, Konstanz

⁶ (Berge 95)

Kommunikationsforen sind damit Instrumente kommunikativer Bildung auf globalen elektronischen Arenen.

3. Design von ComFor für VF-INFOethics

Im Anschluß an die erste UNESCO-Konferenz über Informationsethik im März 1997 in Monte Carlo wurde beschlossen, mit einem elektronischen Kommunikationsforum zu experimentieren, um der UNESCO Erfahrung im Umgang mit dem neuen Medium zu verschaffen und um durch ein elektronisches Forum eine neue und vielleicht höhere Form von Rationalität in der Diskussion über Informationsethik zu erzielen, als es in den standardisierten Austauschformen traditioneller Konferenzen möglich ist. Daß ein elektronisches Forum allemale billiger ist als eine Konferenz mit Teilnehmern aus aller Welt ist ebenfalls ein nicht zu vernachlässigender Faktor.

Die Konstanzer Informationswissenschaft erklärte sich bereit, in konzeptioneller, technischer und Management-/Monitoring-Sicht das Forum zu betreuen. Dazu wurde auch die erforderliche Software entwickelt, die unter dem Namen ComFor (Communication on Forums) in einer funktionsfähigen Version vorliegt.

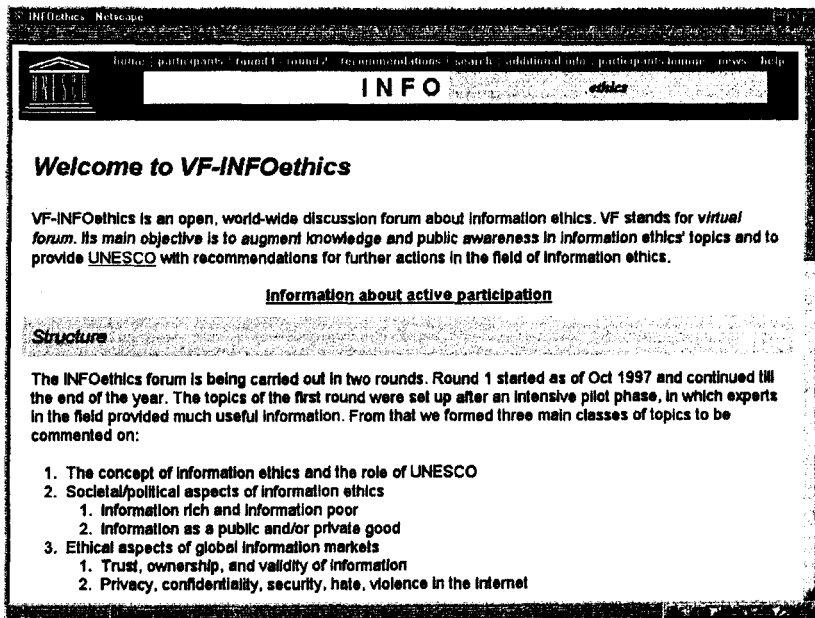


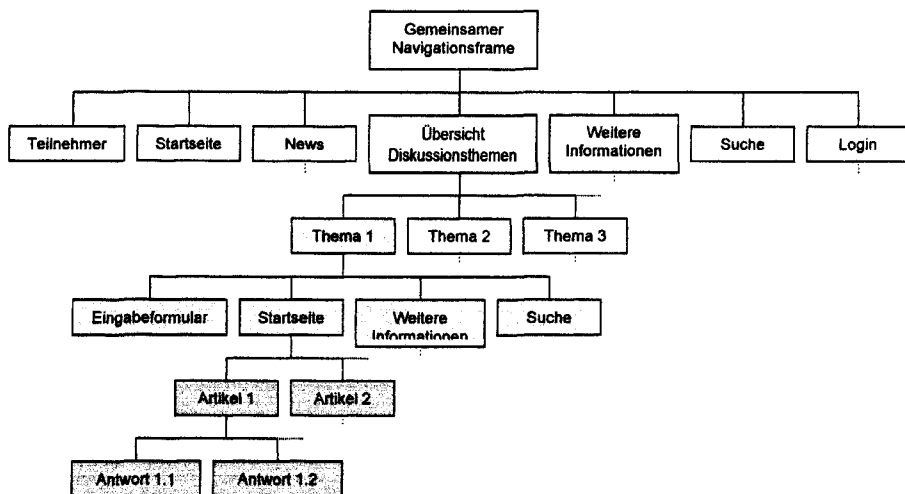
Abb. 1: Startseite von VF-INFOethics

Wesentliche Zielsetzung bei der Entwicklung war, die Software adaptiv und flexibel zu gestalten, so daß sie leicht an verschiedene Verwendungszwecke

und Neuerungen angepaßt werden kann. ComFor soll also offen, adaptiv und skalierbar sein, es soll sich leicht in verschiedene WWW-Anwendungen integrieren und an unterschiedliche Anforderungen anpassen lassen. Aufgrund seiner guten Eignung für WWW-Anwendungen wurde ComFor mit ASP (Active Server Pages) entwickelt.

Implementation des Systems ComFor

Alle Seiten von ComFor werden dynamisch aufgebaut: Entsprechend einigen allgemeinen Anwendungsparametern, der Gruppenzuordnung des Benutzers und dem aktuellen Bearbeitungsstand werden die Informationen und Hyperlinks auf der HTML-Seite ausgewählt. Besucher bekommen z.B. Funktionen, die nur



angemeldeten Teilnehmern zur Verfügung stehen, gar nicht zu sehen.

Abb. 2: Hierarchische Seitenstruktur von VF-INFOethics

ComFor bietet einen guten Navigationskomfort über die hierarchische Ordnungsstruktur (s. Abb. 2). Der Schwerpunkt der Anwendung liegt bei den einzelnen Diskussionen. Es können ein oder mehrere Diskussionsthemen angelegt werden, die wiederum in einer oder mehreren Übersichten zusammengefaßt werden können.

Die Diskussionsthemen an sich bieten einen kurzen einleitenden Kommentar, eine Übersicht der vorhandenen Artikel zum Thema, eine Suche in den Artikeln des Themas und eine Sammlung von themenspezifischen Hyperlinks an. Die Artikelübersicht kann nach verschiedenen Kriterien sortiert werden (dies kann nur vom Administrator eingestellt werden), zum Beispiel rein chronologisch oder in Diskussionsfaden-Ansicht. Dabei werden einzelne Artikel und ihre Antworten als zusammengehörend angezeigt.

Zum Eingeben neuer Artikel steht ein HTML-Eingabeformular zur Verfügung. Nur angemeldete Teilnehmer können neue Artikel schreiben. Die Artikel können mit HTML-Anweisungen formatiert werden; es ist daher eine attraktive Darstellung der Beiträge möglich. Zudem können auch Hyperlinks, Tabellen, Grafiken etc. in die Artikel integriert werden. ComFor arbeitet mit Push-Techniken: beim Eintreffen eines neuen Artikels werden Benachrichtigungen an alle daran interessierten Teilnehmer versendet, die den Autor des Artikels, den Titel und das Diskussionsthema enthalten. Das Versenden der Emails geschieht über eine in die Anwendung integrierte ActiveX-Komponente.

Neben den Diskussionen runden einige weitere Informationen die Anwendung ab, wie z.B. eine Liste aller aktiven Teilnehmer, eine Sammlung von Hyperlinks mit Bezug zum Thema (diese kann von allen Teilnehmern um neue Hyperlinks erweitert werden) oder die Teilnehmer-Lounge, in der die angemeldeten Benutzer ihre Daten einsehen und ändern können.

ComFor ist leicht zu verwalten: Viele administrative Aufgaben können mit Hilfe von WWW-Formularen erledigt werden. Für andere Anpassungen stehen Textschablonen zur Verfügung, zum Beispiel für die verschickten Benachrichtigungs-E-mails. Durch die Offenheit des WWW wird ein großer Teilnehmerkreis erreicht. Das Forum stellt gleichzeitig die Funktionalitäten einer Newsgroup, eines Newsgroup-Archivs und einer Hyperlinksammlung zur Verfügung. Es können also vielfältige Informationen an einem Ort gesammelt und ohne Medienbrüche angezeigt werden. ComFor kann außerdem gut in bestehende WWW-Anwendungen integriert werden; dies ist eine der Stärken von ASP⁷. ComFor ist sehr modular aufgebaut; die einzelnen Funktionalitäten liegen in separaten Dateien vor. Daher können leicht neue Funktionen hinzugefügt oder alte entfernt werden.

Verwendung von Active Server Pages

ASP ist eine Erweiterung des Internet Information Servers von Microsoft. Eine WWW-Anwendung mit ASP besteht prinzipiell aus einer Anzahl von Dateien, die innerhalb eines Stammverzeichnis abgelegt sind, das als virtuelles Verzeichnis im Internet-Dienst-Manager eingetragen ist. Die ComFor-Anwendung enthält Dateien in verschiedensten Formaten: neben den ASP-Scriptdateien zum Beispiel auch HTML- und ASCII-Textdateien, eine Initialisierungsdatei (`Global.asa`), Grafiken im GIF- und JPEG-Format, eine Datenbankdatei sowie eine ActiveX-Komponente.

ASP-Scriptdateien sind prinzipiell HTML-Dateien, in die Anweisungen in einer Programmiersprache⁸ eingefügt wurden. Für das Einbinden dieser Anweisungen stellt ASP einige spezielle HTML-Befehle zur Verfügung. Beim Aufruf einer ASP-Seite führt der Server zunächst die Scriptanweisungen in der Datei aus, bevor er das Ergebnis als reine HTML-Datei an den Client zurückschickt.

⁷ vgl. (Aßfalg 97)

⁸ z.B. VBScript oder Jscript

Jede ASP-Datei stellt quasi eine Mini-Anwendung für sich dar, sie ist zunächst unabhängig von allen anderen Scriptdateien der Anwendung. Für ComFor ist es, wie bei vielen Anwendungen, allerdings erforderlich, Informationen zwischen den verschiedenen Programmteilen weiterzugeben. Dazu bietet ASP mehrere Möglichkeiten an:

- a) mit Hilfe von eingebundenen Dateien (Include-Files), dafür stellt ASP ein `INCLUDE`-Statement zur Verfügung;
- b) über anwendungs- oder sitzungsweit gültige Variablen
- c) mittels Formular- oder HTTP-Übergabeparametern, die von den Scripten erzeugt bzw. ausgewertet werden
- d) durch Konfigurationsdateien (z.B. `Global.asa`)
- e) durch Datenbankeinträge, die von mehreren Dateien geschrieben oder ausgelesen werden.

Die Scriptdateien von ComFor sind stark untereinander vernetzt, da Parameter wie z.B. die Benutzerzuordnung von einem Anwendungsteil zum nächsten weitergegeben werden; dazu werden alle oben genannten Methoden eingesetzt. Ein Beispiel: Alle Scripte von ComFor erhalten über einen HTTP-Parameter eine Sitzungskennung, die zur Verwaltung von Benutzer-Sitzungen benötigt wird. Die Auswertung dieses Parameters geschieht in einer eingebundenen Datei, die von allen Scripten verwendet wird. Die einzelnen Scripte müssen sich also um die Auswertung dieses Parameters nicht selbst kümmern, sie erhalten sie direkt aus einer Methode der eingebundenen Datei übergeben.

Außerdem kann eine Datei mehrere voneinander getrennte Bearbeitungsschritte enthalten. Auf die Weise lassen sich inhaltlich zusammenhängende Funktionen in einer Datei zusammenfassen. Zum Beispiel enthält die Scriptdatei `newComment.asp`

- a) ein Anmeldeformular, das angezeigt wird, falls der Schreibzugriff nicht für jedermann frei ist und der Benutzer sich noch nicht angemeldet hat
- b) die Auswertung dieses Anmeldeformulars
- c) das Eingabeformular zum Schreiben eines Beitrags
- d) eine Beitrags-Vorschau (die Beiträge können nach dem Abschicken nicht mehr geändert werden)
- e) die Auswertung des Beitragsformulars, die das Eintragen des Beitrags in die Datenbank übernimmt.

Jede dieser Funktionen erzeugt eine eigene HTML-Seite und ist mit den anderen Funktionen in der Datei zunächst nicht verbunden. Zur Unterscheidung, welche Funktion gerade gewünscht wird, führt `newComment.asp` daher eine Status-Variable mit, die als Formularparameter übergeben wird. Über Formularparameter werden außerdem die vom Benutzer eingegebenen Informationen weitergereicht.

4. Ergebnisse

VF-INFOethics - einige Ergebnisse

VF-INFOethics ist als virtuelles Diskussionsforum entworfen, das die Funktion der UNESCO als ein weltweites Ideenlaboratorium unterstützen soll. VF-INFOethics soll nicht nur Stellungnahmen und Kommentare der Teilnehmer enthalten, sondern sich zu einer offenen Wissensplattform zu Themen der Informationsethik entwickeln, weitgehend durch Verknüpfungen mit externen, im Internet verfügbaren einschlägigen Materialien.

Die Diskussion in VF-INFOethics wurde in zwei Runden durchgeführt. Die erste Runde dauerte vom Oktober 1997 bis zum Ende des Jahres. Die zweite Runde begann Ende Februar 1998 und wurde im Juni abgeschlossen. Erwartet werden für die UNESCO Empfehlungen und das Einbringen der Ergebnisse in die zweite UNESCO INFOethics Konferenz im Oktober 1998. Am Forum haben ca. 60 Experten aus aller Welt teilgenommen. Bei der Auswahl der Teilnehmer wurde auf den Bezug zur UNESCO sowie auf Ausgewogenheit hinsichtlich der Nationalität geachtet, um ein Übergewicht einzelner Nationalitäten (wie z.B. Nordamerika oder Europa) zu vermeiden. Nur die eingeladenen Experten hatten Schreibrechte, zum Lesen war das Forum web-öffentlich. Es steht im WWW unter <http://www.de3.emb.net/infoethics/> jedermann zur Einsicht zur Verfügung.

Die Diskussion war nach den folgenden Hauptgebieten strukturiert. Diese wurden als Ergebnis einer ausgiebigen Pilotphase formuliert, die noch über Email organisiert war.

- The concept of information ethics
- Information rich and information poor
- Information as a public and/or private good
- Trust, ownership, and validity of information in cyberspace
- Privacy, confidentiality, security, hate, violence in the Internet

Zu jedem Themengebiet wurden ausführliche Fragenlisten und einschlägige Informationen vorgegeben. Für jedes Thema war ein Moderator verantwortlich, dessen Aufgabe darin bestand, die Diskussion zu stimulieren, Zusammenfassungen zu schreiben und am Ende Empfehlungen für die UNESCO zu formulieren (und natürlich noch einmal zur Diskussion zu stellen). Als ein Beispiel für die Diskussionsergebnisse geben wir hier die im Forum erarbeiteten Empfehlungen zum Thema Information rich/poor wieder:

- "To bring net access to poor countries by putting existing resources to sensible use in order to promote the development of global and local information cultures and economies.
- To support the development of a World Information Ethos.

- To support concrete projects in information poor countries in order to create country-specific information centres.
- Public awareness on these matters through virtual forums, publications, and conferences should be promoted.
- To provide permanent, specific, and detailed knowledge of existing information activities in information poor countries.
- UNESCO should promote the rights of non-English-speaking-countries and their economic interests.
- UNESCO should promote topics in information ethics to be included in curricula at all levels.
- Promotion activities through international organizations should be based on grassroots efforts as well as on a decentralized and well coordinated basis.⁹

RichPoor-Topics

Topics: Information rich and information poor

Topics

UNESCO

Comment on this topic

Search in comments

Recommendations

Add topic-suggestions

Information competence:

Will information competence become a basic and necessary requirement or the main prerequisite for success?

Rafael Capurro, the chair of this topic, has formulated a few recommendations as the first results of the discussion. You can comment on them as on every other topic.

- **Ethics of abstraction**
- **Gaps**
- **Information competence**
- **Language**
- **Overabundance of information**

T1 Information poor (Christine Morris, 09.01.99 09:36:56)

T2 Summary 10-27 to 11-09, 1997 (Rafael Capurro, 25.11.97 15:05:19)

T3 Information rich and poor people (Christopher Zielinski, 18.11.97 19:03:42)

- 1 **Re: Information rich and poor people** (Rafael Capurro, 24.11.97 18:49:30)
- 2 **Re: Information rich and poor people** (Michael Gurstein, 21.11.97 17:09:05)

T4 Traditions insufficient (Kristóf Nyíri, 04.11.97 18:49:51)

- 1 ad 2 **Re: Re: Traditions insufficient -global/local** (Kristóf Nyíri, 23.11.97 21:26:06)
- 2 **Re: Traditions insufficient -global/local** (Rainer Kühlen, 09.11.97 01:03:42)
- 3 **Re: Traditions insufficient** (Rafael Capurro, 06.11.97 12:33:57)

T5 Rights and Responsibilities (Rafael Capurro, 04.11.97 12:08:00)

- 1 ad 2 **Re: Re: Again: World Information Ethos** (Rafael Capurro, 24.11.97 18:29:28)
- 2 ad 3 **Re: Again: World Information Ethos** (Rainer Kühlen, 10.11.97 21:33:38)

Abb. 3: Inhaltsverzeichnis der Diskussion zu Information rich/poor

Ingesamt wurde klar herausgearbeitet, daß es für Informationsethik im globalen Maßstab kaum allgemeingültigen Vorgaben und erst recht kaum allgemeingültige Ergebnisse geben kann. Die globalen Netzwerke und

⁹ s. <http://www.de3.emb.net/infoethics/Topics/comm.asp?TopicNr=RichPoor-Recomm&ThreadID=294>

elektronischen Dienste sind der Ort, an dem sich Informationsethik artikuliert und wo sich die Konsense bilden können. Die Teilnehmer waren besonders vorsichtig, westliche ethische Vorstellungen auf die globale Informationsgesellschaft zu übertragen und waren daher besorgt, inwieweit die Konvergenzeffekte im Internet eine unerwünschte Nivellierung von Kulturen und Wertvorstellungen bewirken werden. Ergebnis der Diskussion wird kaum die eine Informationsethik sein, sondern die Entwicklung eines Ethos, durch das die Öffentlichkeit, die Wirtschaft und die politischen Instanzen für die Bedeutung eines ethischen Umgangs mit Information sensibel gemacht werden kann.

Sensibilität für die Rolle von Information (mit allen Implikationen der Produktion, Aufbereitung, Verteilung und Nutzung) ist somit der Hauptgegenstand einer Informationsethik. Die Frage ist in ihrer Bedeutung und den Möglichkeiten ihrer Beantwortung vergleichbar mit Fragen der Ökologie. Es besteht ein Bedarf nach einer weltweiten Informations-Agenda. Ob dafür eigene UN/UNESCO-Institutionen erforderlich sind oder andere Formen des Austauschs gefunden werden, wie elektronische Foren, ist eine offene Frage.

Von allen Spezialfragen stand in der Diskussion die Frage der Informationsgleichheit bzw. -ungleichheit, des Information Gap bzw. das Problem einer informationell balancierten Gesellschaft im Vordergrund. Eine informationell balancierte Gesellschaft ist eine solche, wo auf dem individuellen, sozialen, nationalen, regionalen und globalen Maßstab Informationsdifferenzen (information gaps) im Prinzip beseitigt werden können. Daß dies kein technologisches Problem ist, leuchtet ein, es ist ein politisches, rechtliches, soziales und – unter Einbeziehung all dieser Ebenen - auch ein ethisches. Speziell mit Blick auf die „Gap“-Problematik wurden die folgenden Fragen formuliert, die weiter diskutiert werden:

- "Giving the fact that information is cost-intensive to produce and thus demands a economic basis – who/which institution and which level will be responsible for granting access to appropriate information for the use of local actors?
- How can basic information supply be defined and guaranteed?
- Which models are available to find a fair compromise between market interests in getting information cost investment returned with a reasonable profit and public interest in giving everyone the chance to let their private and public life happen on an information-secure basis?
- Which are appropriate models for such information compromises on a super-national level (overcoming gaps between information rich and information poor countries), a national level (gaps with a information society), and on a individual level (gaps in information competence and education)?"¹⁰

¹⁰ s. <http://www.de3.emb.net/infoethics/report/26.asp>

Wie meistens bei kommunikationsintensiven und komplexen Angelegenheiten hat VF-INFOethics mehr Fragen provoziert als Antworten geben können. Die einzige definitive Antwort bislang ist, daß, nicht nur im Kontext der UNESCO, im Zuge der Globalisierung der Informationsgesellschaft die Diskussion über Informationsethik ein wichtiges Element darstellt und noch bei weiten nicht abgeschlossen ist. Diese Diskussion muß nicht zwangsläufig über elektronische Foren geführt werden, aber diese erweisen sich doch, nach den ersten Erfahrungen, als ein vermutlich dauerhaftes und unverzichtbares Instrument der Konsensbildung bei heterogenen, im Prinzip globalen Interessengruppen und ihren Mitgliedern.

ComFor – einige Ergebnisse

Die bisherigen Erfahrungen mit ComFor sind überwiegend positiv und lassen auf ein großes Potential dieser und ähnlicher Anwendungen schließen. So zeigte sich zum Beispiel, daß die Beiträge in den Diskussionen durchweg qualitativ sehr hochwertig sind.

Wie sich herausstellte, nutzte mehr als die Hälfte der Teilnehmer das Forum lediglich passiv lesend. Dieses Verhalten ist auch von Newsgroups oder Mailinglisten her bekannt. So haben zum Beispiel nur 21 der insgesamt 63 Teilnehmer von VF-INFOethics aktiv an der Diskussion teilgenommen, also etwa 33% der Teilnehmer. Dabei stammen 70% der Beiträge von nur etwa 11% der Teilnehmer.

Eine wesentliche Ursache für die Passivität zahlreicher Teilnehmer liegt wahrscheinlich in der hohen Qualität der Beiträge, die neue Teilnehmer verunsichert und ihren Einstieg in die Diskussion erschwert. Die Antworten auf Beiträge sind also zu wenig spontan; hier wäre vielleicht ein schnelleres und flüchtigeres Medium¹¹ zur Ergänzung der Diskussion sinnvoll. Weiterhin stellte sich heraus, daß die Beteiligung an der Diskussion verbessert werden konnte, indem Moderatoren für die einzelnen Themen bestimmt wurden. Dies wurde für die zweite Runde von VF-INFOethics eingeführt.

Wie sich ebenfalls herausstellte, sind Push-Mechanismen sehr empfehlenswert: die Teilnahme am Forum stieg signifikant, nachdem die Teilnehmer mit Emails über neue Beiträge informiert wurden. Auch scheint Email derzeit noch ein vertrauterer Medium als das WWW zu sein: es wurde wiederholt von Teilnehmern der Wunsch geäußert, per Email auf einen Beitrag antworten zu können.

Ein weiterer wichtiger Ansatzpunkt für die Weiterentwicklung des Systems ist daher die Senkung technischer Eintrittsschwellen in das Forum und die Verbesserung des Bedienungskomforts. Insbesondere die Orientierung und Navigation innerhalb des Forums spielt eine wichtige Rolle.

¹¹ also eine synchrone Diskussion, z.B. über eine Chat-Software

Geplante Weiterentwicklung

Aufgrund der Erfahrungen mit ComFor und VF-INFOethics sowie den anderen ComFor- und KHS-Foren wurde eine Reimplementation der Forum-Software beschlossen: es soll ein neues System entwickelt werden, das die Vorteile beider Foren vereinigt. Das neue System soll wie ComFor leicht skalierbar und unabhängig von teurer Zusatzsoftware sein (so daß man es leicht weitergeben und an andere Anwendungen anpassen kann), dabei aber die Mächtigkeit eines Hypertextsystems wie die KHS-Foren bieten.

Dabei ergeben sich zwei Hauptrichtungen für künftige Erweiterungen: zum einen die Verbesserung der Navigation und Orientierung und des Benutzungskomforts und zum anderen eine Erweiterung und Flexibilisierung der Verwendungsmöglichkeiten. Geplant ist unter anderem eine Anpassung an Benutzerpräferenzen, die Erstellung eines Clients für die grafische Visualisierung der Beitragsstruktur sowie die Ergänzung der asynchron ablaufenden Diskussion um eine synchrone Kommunikationskomponente¹² inklusive einer Terminplanung.

ComFor wird momentan eingesetzt, um die zweite INFOethics-Konferenz der UNESCO mit einem Online-Diskussionsforum zu begleiten¹³.

Literatur

[Aßfalg 97]

R. Aßfalg, U. Goebels, H. Walter: *Internet-Datenbanken*. – Addison-Wesley, Bonn, 1997

[Berge 95]

Zane L. Berge, Mauri P. Collins: *Computer-mediated communication and the online classroom: Overview and perspectives*. (Vol. 1). Cresskill (NJ): Hampton Press, 1995. S. 4

[Fowler 95]

Linda S. Fowler, Daniel D. Wheeler: *Online from the K-12 Classroom*. Aus: Zane L. Berge, Mauri P. Collins: *Computer-mediated communication and the online classroom: Overview and perspectives*. (Vol. 1). Cresskill (NJ): Hampton Press, 1995. S. 83-100

[Höflich 96]

Joachim R. Höflich: *Technisch vermittelte interpersonale Kommunikation: Grundlagen, organisatorische Medienverwendung, Konstitution "elektronischer Gemeinschaften"*. Opladen: Westdeutscher Verlag, 1996.

¹² B. einen Chat

¹³ s. <http://www.de3.emb.net/conference/>

Long-Term Value Adding in an Open Category Network: An Informal Social Approach Towards Relating Conceptual Order Systems on the Internet

Alexander Sigel

German National Research Center for Information Technology,
Schloss Birlinghoven, D-53754 Sankt Augustin, sigel@gmd.de¹

Contents

1. Introduction and Problem Statement
2. Suggested Approach and Underlying Assumptions
3. From Examples Towards a Support Tool
4. Limited Implementation and Experiences with a Predecessor
5. Related Work
6. Discussion
7. References

Summary

Discusses a special case of the problem of comparability and compatibility of conceptual order systems which is of practical relevance when Knowledge Management deals with items from internet sources described with metadata from various order systems. Presents examples from a project which developed an integrated working environment for mediators in the domain of business-to-business contact information brokering. Sketches a tool which could assist in searching/browsing within and across category systems and in gradually capturing and improving knowledge about the interrelations between the order systems. Argues that it is not feasible to start tackling the problem by centralised and authoritative mapping or a stronger formalisation. Suggests instead that users should be better supported in their normal workflow such that they could co-operate on improved mappings. Claims that it would be worthwhile to extend the very limited implementation to test if so much lack of control would really lead to acceptable quality and convergence. (Author)

¹ Now with InformationsZentrum Sozialwissenschaften, Lennéstraße 30, D-53113 Bonn, sigel@bonn.iz-soz.de, where all correspondence should be sent to.

1. Introduction and Problem Statement

This paper asks: What could reasonably be done to improve the current situation in which a plethora of different and ever changing conceptual order systems of mediocre quality is used for categorisation on the internet and, even worse, in which one cannot expect a consensus about their mapping because of many independent players and multiple, potentially conflicting views?

This problem is of practical relevance e.g. when in a Knowledge Management System items from sources are decentrally described with metadata, or when information brokers want to find companies by industries across several internet sources.

During the development of the bizzzyB² prototype of an integrated working environment for information mediators in the domain of business-to-business contact information brokering (EC project COBRA³), it turned out that the old problem of the compatibility of order systems showed up in a new disguise, namely in the specification of appropriate categories in search profiles for web robots which query several networked information sources. Imagine a broker who has to search across yellow pages, business directories, buyers' guides, company listings by region or industry, etc., in order to find companies within a certain industry. As some of these sources actively compete on their categorisation schemata and deny the usage of standard trade classifications and business nomenclatures, and typically no mapping is available between a reference system and one of these idiosyncratic systems, the broker has to use a different schema for almost each source. Together with language barriers this turns out to be quite cumbersome.

2. Suggested Approach and Underlying Assumptions

In contrast to more formal and prescriptive approaches, it is proposed to start with the provision of an open category network environment in which a group of users can gradually build up added value by relating items from various category systems over time. At the same time, this would be an environment for human group filtering. The integration of this component into the users' normal workflow eases motivation problems, as the added value can be shared and reused in the sense of Knowledge Management.

Given the pressing need to immediately support users to expand category sets in order to improve their search specifications, and given an enormous amount of dynamic and highly incompatible order systems, one should not strive for in-depth conceptual analysis, but rather distribute the problem soon to more people, i.e. actively engage the user base, and analyse later.

² bizzzyB (Business-to-business Information Brokering with Categories) is a registered GMD trademark.

³ within the ACTS domain. The acronym stands for Common Open Brokerage Architecture. Short information about the project: <http://zeus.gmd.de/cobra/>

We assume that:

- a useful knowledge structure will evolve because the brokers co-operate in professional working situations on capturing, accumulating and sharing knowledge about categories and relations;
- because of the immediate availability of an evolving knowledge structure, an improvement in metadata descriptions, in profile specifications, or in the improvement process for this knowledge structure itself is expected;
- formal approaches are more appropriate in later stages when a minimal and more stable knowledge base has been set up.

2. From Examples Towards a Support Tool

This section discusses some useful functionalities and illustrates them with selected categories describing *rubber-related products/services* from several category systems, and some of their possible interrelations. Fig. 1 shows a typical set of category systems used to describe company activities or products/services of companies:

a) Internationally Standardised Classifications
<ul style="list-style-type: none"> • ATECO/NACE • SIC/NAICS • Harmonised System • NATO
b) Buyers' Guides / Company Directories
<ul style="list-style-type: none"> • „Wer liefert was?“ (Germany) • Yahoo! Branchenbuch (Classified Company Directory) (Germany)
c) National Yellow Pages
<ul style="list-style-type: none"> • Pagine Gialle (Italy) • Gelbe Seiten (Germany)

Fig. 1: Selected category systems in the domain of company information

Fig. 2 names two examples of machine-readable sources which can be used to provide the system with initial mappings, before users contribute additional mappings:

Official Concordances, intellectually compiled
<ul style="list-style-type: none"> • SIC<->NAICS
Co-occurrences of categories, statistically computed
<ul style="list-style-type: none"> • ATECO<->SIC <p>From a national DB describing companies both with ATECO and SIC codes, from about 55.000 categories co-occurrences above the threshold of 4 co-occurrences were selected</p>

Fig. 2: Selected machine-readable mappings

To give an impression of the difficulty of the mapping task, selected rubber categories from some category systems are shown in Fig. 3a)-h). One can easily see that the construction of category sets is generally not based on firm principles, and that there are no facets.

a) ATECO/NACE

D::Manufacturing

DG::Manufacture of chemicals, chemical products and man-made fibres

24.17::Manufacture of synthetic rubber in primary forms

DH::Manufacture of rubber and plastic products

25::Manufacture of rubber and plastic products

25.1::Manufacture of rubber products

25.11::Manufacture of rubber tyres and tubes

25.12::Retreading and rebuilding of rubber tyres

25.13::Manufacture of other rubber products

b) SIC

3069800::Other Rubber Goods, N.E.C.

3069814::Rubber Coated Fabrics Garment Fabrics

3069841::Rubber Clothing Industrial Rubber Gloves

c) Harmonised System

Section VII, Chapter 40: Rubber and articles thereof

4001::Natural rubber, balata, gutta-percha, guayule, chicle and similar natural gums, in primary forms or in plates, sheets or strip:

4001.10.0000::Natural rubber latex, whether or not prevulcanized

Natural rubber in other forms:

4001.21.0000::Natural rubber in smoked sheets

4001.22.0000::Technically specified natural rubber (TSNR)

4001.29.0000::Other (Natural rubber in other forms, nesoi)

4001.30.0000::Balata, gutta-percha, guayule, chicle and similar natural gums...

4002::Synthetic rubber & factice from oils or mixed with natural rubber in primary form or sheets, strips [...]

4003::Reclaimed rubber in primary forms, plates, sheets, strips [...]

4016::Other vulcanized rubber articles (not hard rubber) floor coverings, mats, erasers, gaskets, washers, seals, inflatables, containers, handles

d) NATO

2630::Tires, Solid and Cushion.

Includes Rubber Track Laying Treads.

3620::Rubber and Plastics Working Machinery.

Includes Plasticators; Presses; Synthetic Rubber Working Machinery;

Thread Extruding Machinery; Vulcanizing Machinery.

Excludes Protective Covering Laminating Presses;

Tire Maintenance and Repair Equipment.

e) Wer liefert was? (134 categories!)

Alterungsschutzmittel für Gummi
Anfahrampenpuffer aus Gummi
Aufvulkanisierbare Gummi-Etiketten [...]
Kautschukschläuche [...]
Gummi-Dämmatten
Gummi-Farbkonzentrate
Gummi-Fussbodenbelag
Gummi-Kunststoff-Metallassemblieren im Lohn
Gummi-Kunststoff-Verbindungen nach Kundenwunsch
Gummi-Metall-Dichtungen [...]
Treibriemen aus Gummi
Verschleissteile für die Gummi-, Kunststoff- und Ölmühlenindustrie
Wandschutzverkleidungen aus Gummi
Zweikomponentenformteile, Gummi-

f) Yahoo!-Branchenbuch

Gummi- und Asbestindustrie
Gummi- und Kunststoffverarbeitungsmaschinen: Hersteller [...]
Gummischläuche: Hersteller
Gummischuhe: Hersteller
Gummiwaren: Handel
Gummiwaren: Hersteller
Kunststoffwaren und Gummiwaren Industrie
Nahrungs-, Genuß-, Chemie-, Gummi- und Kunst-
Technische Gummiwaren: Hersteller

g) Yellow Pages Italy (PagineGialle, Seat)

83386 Bellows metal rubber fabric and plastic
16432 Business agents and representatives chemicals rubber and plastics
76181 Coatings rubber and plastic industrial
89161 Fabrics plasticized and rubberized
65071 Floor rubber linoleum and plastic [...]
92351 Hose rubber [...]
84706 Rubber moulding
41311 Rubber natural and synthetic
41091 Rubber packing and gaskets manufacturers and wholesale
41256 Rubber processing and vulcanization machinery
41201 Rubber products miscellaneous manufacturers and wholesale
83376 Rubber products miscellaneous retail

h) Yellow Pages Germany (Gelbe Seiten, Telekom)

Gummi u. Gummiwaren
Gummiauskleidungen
Gummiartikel
Gummihandschuhe
Gummimetallverbindungen
Gummischläuche
Gummiwalzen

Fig. 3a)-h): Selected rubber-related categories from some category systems
Known mappings from intellectually compiled or statistically computed sources could provide users with initial links (cf. Fig. 4), e.g. to other versions of the

same category system (here: translations, a)), and to suggestions for categories from other category systems (b), c)):

a) Show translations for ATECO 25.1: (intellectual mapping between multiple versions within one category system)	
ATECO	
25.1::Herstellung von Gummiwaren (DE)	
25.1::Fabbricazione di articoli in gomma (IT)	
25.1::Industrie du caoutchouc (FR)	
25.1::Fabricación de productos de caucho (ES)	
b) Show partial mappings for SIC 3069*: (intellectual n:m mapping between two category systems)	
SIC	NAICS
3069::Fabricated Rubber Products, NEC	-
3069:::(partial) Rubberized Fabric	31332::Fabric Coating Mills (partial)
3069:::(partial) Rubber Resilient Floor Covering	326192::Resilient Floor Covering Manufacturing (partial)
3069:::(partial) Other	326299::All Other Rubber
c) Show heuristic suggestions for categories related to ATECO 25.11: (statistical 1:n mapping between two category systems)	
SIC	
15 times co-occurring with 3069000::articoli in gomma n.a.c.	
5 times co-occurring with 2822000::gomme sintetiche (chimiche) e altri elastomeri sintetici vulcanizzabili	
5 times co-occurring with 3011000::pneumatici e camere d'aria	

Fig. 4: Suggestions for ATECO expansion using the initial mappings

Much more interesting are support functions which allow users to contribute added value over time. Then one broker could e.g. look for suggestions of related categories which were made by one of her colleagues during search profile specification. As an example, Fig. 5 depicts what user U5 may see after searching for user-contributed suggestions to extend the category „Gummi-schläuche“ (rubber hoses).

A tool with a uniform interface is needed which supports searching for categories (and their interrelations) as well as convenient browsing in them, regardless of the origin of the categories. Furthermore and most important, the group of co-operating brokers is strategically interested in improving this category network with their own relations and categories. As in different contexts views on the „right“ mapping might differ, and one often cannot judge upon the correctness, the system should allow for conflicting definitions and not exert too much control. In order to be used the system should be rather simple, especially no semantic definition of categories at all is preferred over a logics approach.

User-defined Category	Relation Type	Type of Suggestion	Source of Suggestion	Category System:: Category
Gummischläuche	probably-similar-extension	indirect from co-usage within the same search profile of one user	user U1 date D1	Wer liefert was?:: Kautschukschläuche
	probably-similar-extension	indirect from co-usage within the same search profile of one user	user U2 date D2	Yellow Pages Italy:: 92351::Hose Rubber
	similar-to	direct from user input	user U3 date D3	Yahoo!:: Gummischläuche: Hersteller
	similar-to	direct from user input	user U4 date D4	Yellow Pages Germany:: Gummischläuche

Fig. 5: Example of user-contributed suggestions for category mappings between multiple category systems

4. Limited Implementation and Experiences with a Predecessor

bizzyB is a prototype of an integrated information environment to support human brokerage. The system is implemented with Java, Javascript, Perl, an RDBMS, and HTML. A detailed description of the main ideas of the system and where and how it allows value-adding can be found in (Sigel, Rockenberg & Klemke 1998). (Koenemann & Thomas 1998) provide details about how this system is embedded in the activities of GMD's working group on agent-based information brokering.

In the initial analysis phase of the project, information brokers in the Economic Information Center of the Milan Chambers of Commerce voiced need for a tool to support browsing/searching in a variety of category systems, as they most often searched manually in code books. This requirement was extended to a tool to create and share knowledge about categories and their interrelations. bizzyB's component „WebCatNet" realises this in limited form. The final prototype was evaluated by the brokers.

Unfortunately, the first quite restricted evaluation with the limited prototype does not yet allow us to draw firm conclusions whether a system with the expanded functionality as sketched above would work. This makes us even more curious to see if so much lack of control would really lead to acceptable quality and convergence. Therefore we suggest that it would be worthwhile to extend the implementation and to test the assumptions on a larger scale.

5. Related Work

How to achieve compatibility of order systems has been heavily discussed in the specialty of Knowledge Organisation (cf. e.g. a research seminar devoted to that topic, ISKO *et al.* 1996). (van der Walt 1998) analysed and compared the quality of order systems on the internet and provided simple principle-based suggestions to improve some, e.g. Yahoo! This could be applied to other category systems, too.

In the beginning of the project it was checked if formal approaches (using knowledge representation and logics) could help. One can e.g. achieve semantic integration with ontologies (Vickery 1997, Gruber 1993, Stanford Ontolingua Server o.J.) which may be defined using propositional or description logics.

One source of early inspiration for this work were discussions with Soergel about his open SemWeb proposal (Soergel 1997), as well as with Soergel and Fischer about GMD's formal TerminologyFramework system, implemented in Smalltalk (Fischer 1998, Fischer, Möhr & Rostek 1996, Fischer 1995, Rostek & Fischer 1988).

The idea of learning an evolving conceptual structure from user queries over time can already be found in the work of (Jüttner & Güntzer 1988). They tried to learn thesaurus relations from user queries to an OPAC, and employed explicit feedback to ensure quality.

6. Discussion

It is not claimed that an interactive environment for co-operatively relating heterogeneous categories was something very new. But new is the situation where a proliferation of many intentionally independent categorisation schemata is used to organise internet data, and that they are so dynamic that standardisation and central efforts inevitably fail. Furthermore, a long-during intellectual effort to formalise the categories' semantics is out of reach until contributions in a social setting will have provided some baseline structure. This paper claims that the combination of the following issues has not been encountered before:

- The application to a large number of business categories from several category systems (which is of rising importance to e-commerce);
- The focus on value-adding across several category systems;
- The highly interactive groupwork environment feasible via network and its effects on knowledge management;
- The focus on order systems on the internet (while compatibility is usually discussed for rather stable universal order systems, cf. e.g. (Dahlberg 1996)), and
- An open social, not a standardising or formal approach.

During the project it became clear that appropriate support for category structures - although not originally planned - was a core requirement, and that information science offers interesting contributions in this area.

With only preliminary work carried out, it remains an open question whether in practice formal approaches or informal collaborative environments are more appropriate to improve mappings between conceptual systems. In the author's view, after a period of value-adding with the latter type of systems, the results could be used within the former approach. Formal approaches were not employed because the categories downloaded from the web sites were not even associated with definitions but are only text strings, and the data volume was rather high. However, support for category search and browsing was immediately in need, and there were users available who could contribute to the improvement of the resulting network. These users could not handle logical definitions but preferred an informal environment.

Advantages of a social approach to the compatibility problem are:

- Less time is needed to achieve a useful structure because no consensus is necessary.
- Subject specialists can immediately bring in the vocabulary they work with and are more directly involved.
- Even conflicting views are allowed.
- Already in the beginning phase of Knowledge Management (in which structures are still sought after) added value can be created.

Disadvantages are:

- As there is almost no formal semantics, one cannot support inferences or e.g. consistency checks.
- There remain doubts about the user motivation, the reward system and the critical mass of „good“ relations.
- It is not guaranteed but only expected that the structure will converge or be of sufficient quality.
- There are not yet enough experiences with so much lack of control.

7. References

[Dahlberg 1996]

Dahlberg, Ingetraut: *Library Catalogs in the Internet: Switching for Future Subject Access*, Green, R. (Ed.): Knowledge Organization and Change, Proc. 4th Int. ISKO Conf., Washington, DC, 1996, Frankfurt/M.: INDEKS, 1996, p. 155-164.

[Fischer 1998]

Fischer, Dietrich H.: *From Thesauri towards Ontologies?*, Structures and Relations in Knowledge Organization, Proc. 5th Int. ISKO Conf., Lille, France, 1998, Würzburg: Ergon, 1998, p. 18-30.

[Fischer, Möhr & Rostek 1996]

Fischer, Dietrich H., Möhr, Wiebke & Rostek, Lothar: *A modular, object-oriented and generic approach for building terminology maintenance systems*, Proc. TKE'96, Frankfurt/M.: Indeks, 1996, p. 245-258.

[Fischer 1995]

Fischer, Dietrich H.: *An executable specification language for thesauri and their operational semantics* (Abstract), Learning and Knowledge: Structural aspects, quantitative methods and new applications, Proc. 19th Annual Conf. of the Gesellschaft für Klassifikation, Univ. Basel, 1995.

[Gruber 1993]

Gruber, T.R.: *Toward principles for the design of ontologies used for knowledge sharing*, Guarino, N. & Poli, R. (eds.): *Formal Ontology in Conceptual Analysis and Knowledge Representation*, Kluwer, 1993 (= Stanford Knowledge Systems Laboratory Report KSL-93-04). Also online: <<http://ksl-web.stanford.edu/knowledge-sharing/papers/onto-design.ps>>

[ISKO et al. 1996]

International Society for Knowledge Organisation / Polish Librarian Association / Society for Professional Information: *Compatibility and Integration of Order Systems. Research Seminar Proceedings of the TIP/ISKO Meeting*, Warsaw, Poland, 1995, Warszawa: Wydaw. SBP, 1996.

[Jüttner & Güntzer 1988]

Jüttner, Gerald & Güntzer, Ulrich: *Methoden der künstlichen Intelligenz für Information Retrieval*, München: K.G. Saur, 1988.

[Koenemann & Thomas 1998]

Koenemann, Jürgen & Thomas, Christoph G.: *Agent-Supported Information Brokering*, KI - Künstliche Intelligenz 3/98 (=Themenheft Intelligente Informationsagenten).

[Rostek & Fischer 1988]

Rostek, Lothar & Fischer, Dietrich H.: *Objektorientierte Modellierung eines Thesaurus auf der Basis eines Frame-Systems mit graphischer Benutzerschnittstelle*, Nachrichten für Dokumentation 39, 1988, p. 217-226.

[Sigel, Rockenberg & Klemke 1998]

Sigel, Alexander, Rockenberg, Anja & Klemke, Roland: *Brokering von Firmeninformationen mit bizzzyB*, Proc. Symposium Information Brokering, Leipzig, 1998, Wiesbaden: Vieweg & Sohn, 1998, p. 129-137.

[Soergel 1997]

Soergel, Dagobert: *SemWeb: Proposal for an open, multifunctional, multilingual system for integrated access to knowledge about concepts and terminology: Exploration and development of the concept*, expanded version (1997) of a paper published in *Advances in Knowledge Organization*, Vol. 5, Frankfurt/M.: Indeks, 1996, p. 165-173.

[Stanford Ontolingua Server o.J.]

<<http://ontolingua.stanford.edu/frame-editor>>

[van der Walt 1998]

van der Walt, Marthinus S.: *The Structure of Classification Schemes Used in Internet Search Engines*, Structures and Relations in Knowledge Organization, Proc. 5th Int. ISKO Conf., Lille, France, 1998, Würzburg: Ergon, 1998, p. 379-387.

[Vickery 1997]

Vickery, B.C.: *Ontologies*, Journal of Information Science, 23(4), 1997, p. 277-286.

Teddy Bears and Perfumes: Advertising Information- Knowledge Management at its' best!

M. Burke

**Department of Information and Communications
Manchester Metropolitan University**

Summary

This paper will examine the issues surrounding marketing in today's business orientated climate and the role of knowledge managers as providers of information.

1. Introduction

According to the Institute of Marketing, marketing is the management process which identifies, anticipates and supplies customer requirements efficiently and profitably. Although at first sight this definition appears to confine the term to the activities of a marketing department, it actually implies something much wider. The management process refers not only to the internal organisation of marketing, but also to other elements implicitly present which include market research (identifying and anticipating); correct design and delivery of product service (customer requirements) and satisfactory pricing policy (profitability). All of these functions can also be categorised as the very essence of what information specialists now term knowledge management. But how can knowledge management be properly defined? One of the most succinct definitions is that by Hoyt (1) - "knowledge management in it's most basic form is made up of the systems that connect people to people and people to information". He goes on to say that a key aspect of knowledge management is understanding how organisations, markets, and individuals create and manage value. The corresponding definition of organisational knowledge is said to be the combination of collective and personal information in context – in essence, information with value.(2)

Knowledge management can therefore be seen to be a combination of all the different types of collective information within the organisation. If marketing information is the essential key to success of that organisation - then the skills

that are necessary in order to achieve that level of success must be those of a finely honed knowledge management team.

The modern view of marketing on the other hand is that it is not just a separate function of a specific department of an organisation but that a marketing outlook should be taken by all employees of the organisation if the business is to prosper. A classic analysis of marketing breaks it down into four main elements, namely product, price, place and promotion (the marketing mix). Successful marketing means not only that these four elements must be individually correctly positioned, but that they should be consistent with each other in order to produce a synergistic effect. Thus, all types of products from expensive perfumes to children's teddy bears are correctly designed for a chosen type of customer and must be priced accordingly, distributed at the right places and suitably advertised and promoted. However one looks at it, correct marketing - and the management of this knowledge throughout the organisation - must be the top-rank objective of any business, and all other activities contribute to this aim. In this sense, marketing is the concept that unifies all activities of the firm. This paper presents an examination of how closely intertwined marketing, information and the management of that information in the organisation really is - i.e. the management of knowledge within all parts of the organisation. This is achieved by an examination of the different types of markets; by presentation of those factors affecting both markets and knowledge management and a discussion on the merits of linking together these two areas.

2. Market classifications

In it's narrowest sense, a market is made up of the types and numbers of customers for a given product or service. However, in a wider view, a market may be considered to be the entire exchange system of producers, wholesalers, retailers and customers dealing in the product or service. Markets can be classified in different ways, as follows.

2.1 Consumer and Industrial Markets

It is sometimes helpful to distinguish between consumer markets and industrial (or business-to business markets). In the case of mass-produced branded consumer goods, for example, the manufacturer has thousands of unknown buyers who purchase the product through retail outlets: in the case of industrial equipment and supplies, there may be only a handful of firms who buy the product and they are well known to the manufacturer, who may cultivate their business carefully on an individual basis. The terms of the goods and services which are sold to other business are quite different from those applying to consumer markets, and industrial marketing can be seen as a separate branch of the marketing profession.

2.2 Goods and Services

An obvious distinction is between goods and services: and within goods, between consumable and durables. Another concept commonly encountered is that of fast moving consumer goods, sometimes abbreviated to FMCG,

meaning food, drinks, chocolates, household cleaning material and similar items which turn over very quickly. One of the reasons for introducing these terms is that sources of information are often structured in corresponding ways. Thus, some market reports or compilations of statistics may concentrate on, or be confined to, some particular class of goods or services.

2.3 Geography

A further analysis of markets can be made on a geographical basis, one distinction being between home and export markets. A moment's thought is sufficient to realise the host of additional problems presented by export markets - as opposed to home markets - foreign languages, overseas standards and regulations, transport and distribution problems different patterns of market behaviour and so on. Other geographical analysis of markets might distinguish between local markets, regional markets, national markets, European markets - and the international markets. The European market is worth a special mention here as this market is undergoing many changes, such as new countries joining the European Union; European countries who are presently pursuing accession strategies in order to join the EU and former Communist countries who are now able to trade more freely with other countries. All these reasons make the European market a rapidly changing and diverse playground for many new and established companies.

3. Factors affecting markets

The behaviour and structure of markets are influenced by a complex network of facts, some of them represented by the above considerations of geography and the numbers of buyers and seller. Other important factors include the following:

3.1 Transport facilities and costs

Good or bad transport facilities whether by road, rail, air or sea are powerful determinants and poor facilities may prevent a market from existing at all. The markets for perishable foods have been greatly extended in recent times by a fast motorway network and the development of refrigerated container vehicles.

3.2 Trade tariffs

Governments impose tariffs (import duties) and quota restrictions on quantities permitted to be imported, or even outright bans on certain goods. Tariff barriers between members of the European Union have long since disappeared, but the UK imposes import duties on many products imported from countries outside the Union.

3.3 Consumer groups and the "green" movement

Manufacturers and retailers are increasingly being affected by the environmental and consumer lobbies. Many organisations are now having to pay some attention to environmental issues and this has manifested itself in various ways such as less packaging on goods, information regarding the

ingredients of labelled products and the availability of recycling depots at larger superstores.

3.4 Controls on advertising

In some countries markets such as the tobacco and cigarette market are not free to promote their goods as they wish, and the controls exerted on advertising, both statutory and voluntary, have an effect on consumer behaviour

3.5 The nature of the consumer

The level of demand for consumer goods is influenced by a variety of facts such as the amount of personal disposable income available for the purchase of optional items. Individual tastes in goods vary, and may be affected by a number of psychological and other variables which are difficult to study and to take account of in marketing. Tastes can, of course, be influenced by group behaviour, as in the case of clothes fashions and manufacturers themselves can create fashions or tastes in some sectors by clever promotion.

3.6 Prices

Price influences choice in complex ways. Thus for example, people may buy more fresh strawberries when the price falls than when they first appear on the market at a very high price. But, however low the price may fall, customers will not necessarily go on buying more - they may become surfeited with strawberries or some other attraction arises, so that the marginal utility of an extra strawberry is seen to be too low. Prices also relate to substitute and complementary goods.

3.7 General and Social Economic factors

Social and economic factors include changes in the size and composition of the population, which will have some effect on the markets for products and services aimed at these groups. Similarly, uneven distribution of income and wealth in society will have an effect on markets. Factors affecting supply include such things as the prices and supply of raw materials, the cost of labour, energy and other operating costs, the efficiency or otherwise of the technology in use, and so on. Also important to consider are political issues of the moment - which may affect a particular industry - and such things as the economy stability - or instability - of the suppliers country.

4. Factors affecting knowledge management

Clearly the factors outlined above have implications for the supply of information. Some of this will come from a company's own in-house files, such as analyses of sales of the firm's products, to indicate likely or regional trends. Much information about the market generally, social and economic trends, the fortunes of competitive brands and products, relevant legislation and regulations and so on, will be available from published sources. Finally in the

absence of published data, a firm may set out itself to research buying behaviour and other aspects of its markets.

Whether research focuses on the marketing of a specific product or on understanding and monitoring of a market generally, information in the categories mentioned above is likely to be required. What factors then will affect the successful knowledge management of an organisation?

4.1 Group dynamics and politics

All organisations have what is termed "office politics" - especially in very large organisations. The free flow of information - both horizontal and vertical flows tend to be affected by the group dynamics present in the organisation. This is sometimes one of the knowledge managers most difficult problems - if the team need full access to most information in order to provide an excellent service - then a hampering of the free flow of information can cause grave difficulties.

4.2 Technology

The technology which is available to the team is crucial to the success of an overall knowledge management policy. The team need to be able to keep in touch with what is happening in different departments in the organisation - they need access to email, faxes, mobile phones, -ideally they need the latest equipment such as laptops and state of the art computers. Today's knowledge managers do not spend all their time in the information unit or library - but are out in the organisation talking to other members of staff and keeping abreast of new issues, policy changes, trends and new products.

4.3 Professional skills

Good quality, well trained, alert and innovative staff who have a sense of business acumen are indeed an asset to the organisation. Good staff also expect - and deserve - adequate remuneration which, unfortunately in some organisations is often seen as a "cost" which can be glossed over.

Also worth a mention here is the cost of continuing professional development - staff must be encouraged to develop themselves and their careers - but again this is often seen as a disposable cost by the employing organisations.

4.4 Facilities

The facilities which are provided for the knowledge management team will also have a bearing not only on success but also on the efficiency and effectiveness of the team. Adequate staffing resources, adequate finances to buy material and staff training are all a cost to the organisation but it is essential that this is seen to be vital to the success of the marketing not only of the products or service - but ultimately of the organisation itself.

4.5 Organisational structure and communication

The structure of organisations inevitably play some part in the general culture and atmosphere which predominates in the work place. This in turn affects communication. The shape of the organisations can also affect the types of communication eg. a flat structure where the organisation is very "open" or

Weber's typical bureaucratic structure which has a tendency to rely on only on rigid vertical lines of communication. It is not the remit of this article to detail all the variables of communication in organisations but merely to point out that all of these facets play some part in influencing the success of the knowledge management team who need to be able to gain information easily from all aspects of the firm.

These and no doubt a host of other factors all contribute to the success of knowledge management working at its best in organisations.

5. Discussion

Cropley(3) stated that "knowledge managers, must ready themselves for dramatic changes and position all their services and activities carefully if they are to gain the full advantages of the technique." We are presently in a time of such rapid change and no doubt the next few years will see increasing changes - that all business and marketing organisations will struggle to keep up with the changes. Organisations more and more will come to rely on their knowledge managers for helping them sift through the streams of data and vast realms of paper which flow in and out of buildings. It is worth considering the context in which business organisation are operating in the final months of the twentieth century. Since the 1980's a further decade of technical advances and deregulation combined with increased trade freedom in much of the world have had a profound impact on the types and organisations of business enterprise. *Knowledge managers must therefore prove the worth of their services in providing information that can, for example, lead to better decision making; move projects from one stage to the next; facilitate improved relationships with clients; exploit new business opportunities and save time and money* (4). The knowledge managers specialist training in the discriminatory techniques of the selection, acquisition, abstracting, cataloguing and classification, indexing and packaging of information resources should be promoted. This can be achieved by establishing the basic principles of integrated information services provision.

6. Conclusion

This paper deals with marketing and the role of knowledge managers in today's market place. Products must not only be correctly promoted and priced, but also made available in the right places. To great extent the pattern of distribution will be determined by the nature of the product. At the same time, the existence of certain types of outlet has encouraged the development of new products and different types of packaging. For Central and Eastern European countries there has never been a more important time for the knowledge manager to assist organisations with this kind of business information. Marketing therefore provides knowledge managers with an ability to use their expertise in advertising their main product - information. Information is the essential and vital tool on which all marketing principles are formed. From Teddy Bears to Perfume, no matter what the product, the management of

knowledge from within and around the organisation is to be one of the ways in which the skills of trained information specialists are to be used now and in the future.

References

1. Hoyt, B. *Definitions of knowledge management* Knowledge Management News, May 1988, p4-6
2. As above
3. Cropley, J. *Knowledge management : a dilemma*. Business Information Review, 15 (1) March, 1998, pp27-34
4. Burke, M. and Hall, H. *Navigating Business Information : a practical guide for information managers*, Library Association, 1998

Vague Transformations in Information Retrieval

Thomas Mandl

Social Science Information Centre, Bonn
Lennéstrasse 30; 53113 Bonn; ma@bonn.iz-soz.de

Content

- 1 Introduction
- 2 Heterogeneous Datasources in IR
 - 2.1 The Layer Model
 - 2.2 Text-Fact-Integration
- 3 Transformation Methods
 - 3.1 Statistical Methods
 - 3.2 Transformation Network
 - 3.3 The COSIMIR-Model for Transformations
 - 3.3.1 Neural Networks in IR
 - 3.3.2 The COSIMIR Model
 - 3.3.3 Adaption of COSIMIR for Heterogeneous IR
- 4 Conclusions
- References

Abstract

Information Retrieval (IR) is often confronted with different and heterogeneous collections of multilingual and multimedial content. Transformations are necessary in order to map between different representation schemes which are the result of different content analysis methods and different terminologies. Vague methods which have been successfully applied to IR seem especially suited to tackle the vague nature of such transformations. Statistical methods are reviewed and two experimental systems based on neural networks are presented.

1 Introduction

The advent of multimedia and the ever growing number of database providers shape today's reality in Information Retrieval. Users face an overwhelming

number of resources in the internet and other information markets. In many cases, users want to query a wide range of resources as it is not known in the beginning of an information process whether the information need can be satisfied by text documents, factual data or multimedia objects. However, information resources and their interfaces are often organized by formal criteria, such as numerical data, text documents and collections of internet links. End users as well as information brokers prefer one integrated interface over many interfaces. Alta Vista and Yahoo are popular examples for systems integrating many internet sources.

The heterogeneous nature of the data is a great challenge for IR which works on representations of the indexed objects. The representation of text documents consists most times of keywords or index terms which are chosen by a human indexer or derived by an automatic indexing algorithm. Multimedia objects like pictures or video clips may also be represented by keywords and new representation schemes based on their specific features such as color distribution or video scene analysis have been developed. Mappings between keywords and specific multimedia representations nevertheless remain difficult. In addition, problems can occur when integrating different text collections which have been indexed by different methods or using different thesauri. For example, the meaning of the same word in two fields can differ considerably. Aligning words found by automatic indexing with corresponding terms assigned by human indexers will likely result in inconsistencies and poor retrieval performance. Currently, these differences are widely neglected. Although the Internet proves that searches over heterogeneous documents are possible, the quality of the results also shows that many of the conceptual problems of data integration remain yet to be solved. The next chapter further elaborates on the problems arising from various types of heterogeneity in IR. The layer model deals with text collections of varying data quality and different content analysis methods. The text-fact integration in ELVIRA points towards the integration of multimodal data types.

Chapter three reviews vague methods which can be employed to transform between different representation schemes. The family of statistical methods has already been utilized in real world IR problems and seems therefore very appropriate. However, their computational abilities and their learning capabilities suggest neural networks as a candidate for transformations. Two backpropagation networks allowing integrated IR are discussed. The transformation network already applied for an experimental system has a high plausibility. The adapted COSIMIR (COgnitive SIMilarity learning in Information Retrieval) model is interesting as it does not need an explicit transformation step, however, so far it has only succeed when used with small data sets.

2 Heterogeneous Datasources in IR

A classical example for retrieval from heterogeneous document collections is Multilingual Information Retrieval. It allows users to formulate their query in one language and to retrieve documents in various others (cf. Hull/Grefenstette

1996). The multilingual corpus can be seen as a meta-database consisting of several monolingual datapools. The most interesting case is the retrieval of a document in a language other than the query language. In this case, the query or the documents must be transformed into a representation in the other language. After that, techniques from monolingual IR can be applied.

The need for transformations in Multilingual Information Retrieval is evident as the vocabulary between languages is basically distinct. In areas where the vocabulary partly overlaps the heterogeneity is often ignored. Two examples for such problems are introduced in the remainder of this chapter. Text-Fact-Integration provides an example for multimodality whereas the layer model integrates datapools of different quality.

2.1 The Layer Model

The layer model (Krause 1996) has been developed for information service institutions. It suggests deregulation to overcome centralistic structures and its political dimension calls for a new information-science model of information provision. Its key idea is the replacement of a monolithic database by a layer structure. The layer model allows various levels of relevance and content analysis where norms and quality are not imposed, but rather coordinated and administered. That means, an information provider does not need to fulfill all the requirements (use of a certain thesaurus, data quality) to be included in the monolithic database of an information service center. Rather, any information provider can choose to be included in an outer layer where requirements are relaxed. The innermost layer contains the nucleus of the most relevant documents where content analysis is deep and of high quality. Towards the outer layers, relevance and quality of content analysis continually decrease. To allow flexible querying over several layers, transformations between different content analysis schemes (e.g. thesauri) need to be established.

The layer model provides a framework for the common situation of many information centers which have performed intellectual indexing for a corpus of controlled quality. In many cases, users are willing to relax their requirements on quality of the data when they can query a larger document base. Therefore, information centers include new data pools, which have been either indexed by other institutions using different thesauri or which have not been indexed yet. Often, these new data sources cannot be subjected to the same deep content analysis due to economical or political reasons. As a result, automatic indexing is used. Nonetheless, the intellectual work for thesaurus development and indexing should not be lost, but rather included in the retrieval process.

This leads to non trivial problems as the same term used in different thesauri can have different meanings. In addition, a term found in many documents by an automatic indexing algorithm has a different value for retrieving documents than the same term given to few documents by a human indexer.

2.2 Text-Fact-Integration

The integration of textual and factual data (statistical time series data and other numerical data in tables) is the focus of the project ELVIRA¹ (Scheinost et al. 1998; Krause et al. 1998). ELVIRA is an online information system for time series distributed by three German industrial associations to their member companies. Empirical studies with users have shown that the market researchers in companies also want to retrieve text documents to satisfy their information needs. A prototype of ELVIRA including text retrieval as well as time series has been implemented. To achieve a user friendly system, the integration needs to be further developed and should allow the retrieval of both data types using one query.

First empirical studies have shown that the indexing vocabularies of texts and facts differ greatly due to the different ways of content analysis. The text documents are indexed automatically resulting in a weighted document-term-matrix, whereas the time series are indexed intellectually using official hierarchical nomenclatures. For example, the controlled nomenclatures for the time series contain only one term for a product whereas the text collection typically contains many synonyms for the same product. Often, the nomenclatures contain an official term not commonly used in everyday language. Therefore, transformations from the indexing terms of one data type to another are necessary.

3 Transformation Methods

The IR process in an integrated system which takes into account the different representation schemes must include a transformation step proceeding the standard IR process. The transformation step needs to assure that both query and document are represented within the same scheme to enable standard IR processes. The transformation of the query into the representation scheme of the according document collection seems to be the most efficient way. However, this method raises the question of whether a user really uses a specific representation scheme when formulating his query. This is indeed the case when he is either an expert in the use of one thesaurus and its terminology or when the interface solely allows thesaurus based access, like the time-series tool of ELVIRA. It is often the case however, that the user will use everyday language and the system transforms his terms into specific thesaurus terms used in the collections in question.

Which methods can be used to implement transformations? One approach is either to construct a single thesaurus for the field or develop concordances between existing thesauri. Multilingual IR often utilizes these approaches. The

¹The project ELVIRA has been funded by the German Ministry of Economy, grant no. II C7-003060/10 and IV C2-003060/22.

construction of a single thesaurus is expensive and is not advisable for each field as each existing thesaurus represents a viewpoint which can be appropriate for certain user intentions. Concordances are basically rule-based systems where a rule transforms one term into its synonym in the other thesauri. Although this seems a straightforward method, it bears several problems besides the cost factor. For instance, the mapping may depend on the context and rules may be more complex than simple synonym to synonym transformations. Many thesauri have been developed for distinct purposes and are differ so much that simple mappings are rarely successful and that is especially the case for the large vocabulary derived from automatic free text indexing in which intellectually constructed concordances are not feasible.

In the context of ELVIRA, a concordance between the old nomenclature for statistics of industrial production which was valid until 1994 and the new nomenclature introduced in the European Union in 1995 was developed. Groups of products were broken up and the single products were distributed over distant positions in the new nomenclature where they formed new groups with different products.

Although concordances may work well in some areas, empirical investigations suggest that transformations are a fundamentally vague problem. Therefore, vague information processing methods already successfully applied to IR should be tested in order to implement transformations.

The following paragraphs discuss three vague methods for transformations. The first is based on statistics and has been used for real world multilingual IR (Sheridan/Ballerini 1996). The second and the third approaches are based on neural networks and have so far been only tested with small experimental data sets.

3.1 Statistical Methods

In cases where manual construction of concordances is difficult e.g. between a thesaurus and the indexing vocabulary of automatic indexing, the automatic construction of concordances becomes desirable. Several statistical approaches designed to achieve this goal have been proposed.

Biebricher et al. 1988 developed the system AIR/PHYS to increase the quality of manual indexing by automatically suggesting appropriate indexing terms. Some already manually indexed documents were additionally analyzed by an automatic indexing algorithm to calculate association factors between free text terms t and thesaurus descriptors d :

$$z(t, d) = \frac{h(t, d)}{f(t)} \quad (\text{Biebricher et al. 1999:334})$$

$f(t)$ number of documents containing t

$h(t, d)$ number of documents from $f(t)$ indexed with d

For the thesaurus of 22 000 terms, 800 000 association factors were calculated, of which the 350 000 most important were included in the system. The transfer function was enriched with 50 000 USE and 170 000 BROADER TERM relations from the thesaurus. Biebricher et al. 1988 report that the indexers could use many of the suggestions of AIR/PHYS.

Statistical methods with their inherent vagueness have also been utilized for multilingual IR. Sheridan/Ballerini 1996 developed a system similar to the first one where they changed the role of documents and terms, such that each term is represented by a feature-vector in the document space. A similarity matrix for all terms can be calculated. The knowledge for the transformation is derived using a corpus with comparable documents. Sheridan/Ballerini 1996 used news stories which were available in two languages. Thus, the document space is the same for each language and each term can consequently be transformed into its document vector. Since the same documents exist in the other corpus, the document vector can be used to calculate the corresponding term vector for the second language.

As a result, query terms in one language are not translated into a single word using a dictionary, but rather transformed into a weighted vector of terms in the second language representing the use of the query term in the document collection (see figure 1).

Sheridan/Ballerini 1996 tested this method with a collection of 93 000 news stories available in German and Italian. As expected the retrieval results were outperformed by the best mono-lingual experiment, but the overall quality was nevertheless satisfying.

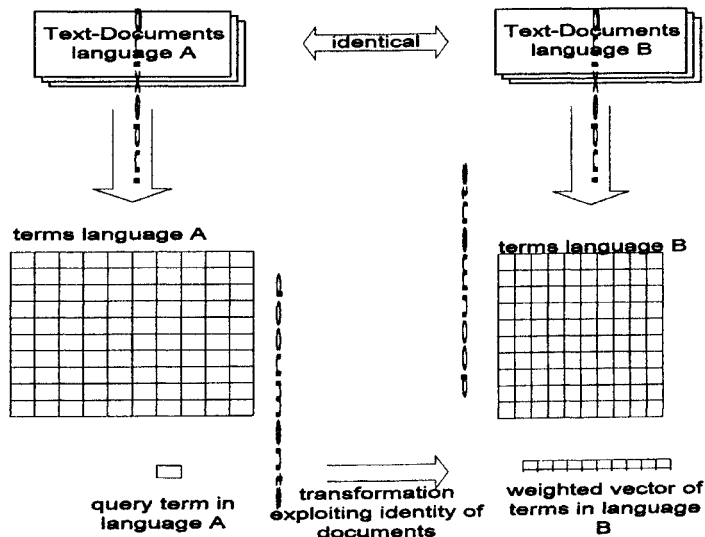


Figure 1: Transformation for multilingual IR (cf. Sheridan/Ballerini 1996)

3.2 Transformation Network

Although statistical methods are well established for vague transformations it is not clear whether they can implement every transformation. The unknown transformation function can be rather complex.

Statistical transformations or concordances implement basically a large rule based system. Therefore, only functions which can be expressed in such a system can be implemented. As the formal nature of the transformation function is unknown, it is advisable to choose a powerful mechanism for its implementation.

Neural networks and in particular the multilayer backpropagation network can learn complex functions such as non linear separable problems (cf. Zell 1994:99). Moreover, Smolensky 1988 argues that connectionist systems such as neural networks including hidden units realize the subsymbolic paradigm of computing and can implement an intuitive processor with capabilities beyond those of symbolic processing. Their knowledge is not derived from rules but from examples which are used as training data. Thus, neural networks can approximate intuitive expert knowledge which cannot be formalized in a rule based system.

All calculations within neural networks are local. Each unit gathers its input from incoming connections and calculates its own activation and output. Signals travel along connections and are modified according to the connection strengths. A typical backpropagation network consists of an input and an output layer and of one or more hidden layers. The input vector is propagated into the network which calculates the output vector. In a training phase, examples for the desired mapping are presented enabling the network to compare the actual output and the desired output. The error is used to tune the connection strengths of the network such that it is closer to the teaching output and that it *can better approximate the desired function*. After training the network, its generalization capabilities need to be tested with another data set (for details cf. Zell 1994).

Crestani/Rijsbergen 1997 present a backpropagation network for the transformation between different queries in which the representation schemes are equivalent. The same architecture can be modified for general transformations between different representation schemes. The goal of Crestani/Rijsbergen 1997 was to transform actual user queries into queries which achieved better results for the users' information needs. A set of query pairs is necessary in order to train the network, each consisting of the original user query and of an improved one. The improved query was constructed using relevance feedback information. The query transformation reported in Crestani/Rijsbergen 1997 led to queries with similar performance as the original query, although the adapted queries were considerably different from the original ones. This shows that the network is not merely an implementation of query expansion where additional terms are added to a query. The Cranfield I collection served as a testbed. Because of its limited size, more testing is necessary in order to show the feasibility of this approach. The following figure shows an adaption of the network for general transformations between two

representations. This architecture is currently used to test mappings between the thesaurus and the classification for the documents of the Social Science Information Centre.

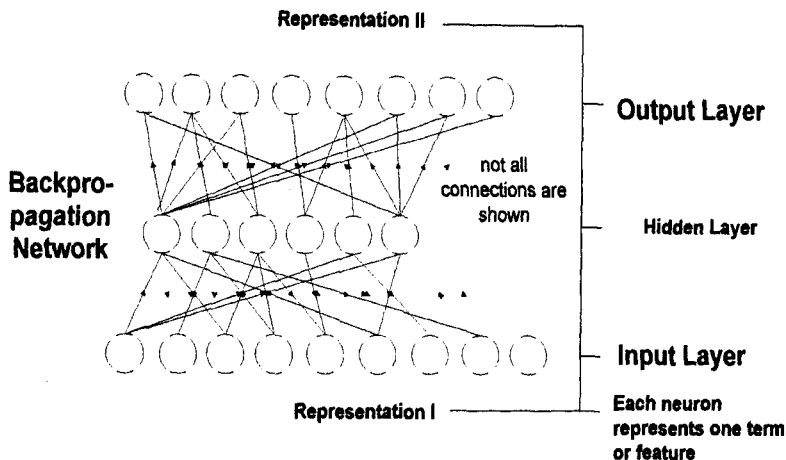


Figure 2: Transformation-Network

The transformation network depends on the same training data as the statistical method for multilingual IR, a corpus of comparable documents in the same feature space.

3.3 The COSIMIR-Model for Transformations

In the same way as the transformation network, the COSIMIR model is based on backpropagation. Currently, most IR-systems consisting of neural networks use Hopfield networks or Self Organizing Maps. These networks do not include hidden units and, therefore, they cannot implement an intuitive processor. COSIMIR is a simple backpropagation network which takes document and query representation as input vector and the relevance between them as the only output unit. So far COSIMIR has only been tested with small problems; however, it can be easily modified to allow retrieval over different representation schemes without an explicit transformation step.

3.3.1 Neural Networks in IR

Neural networks have been applied to IR in many experimental systems (for overviews cf. Doszkocs et al. 1990, Chen 1995, Mandl 1998a). Most systems use either Self Organizing Maps (Kohonen networks) or the so called spreading activation networks (Hopfield networks with layer structure). The latter have

been successfully introduced at TREC (Text Retrieval Conference) where their performance is comparable to that of standard algorithms (Mercure, cf. Boughamen/Soule-Dupuy 1997; PIRCS, cf. Kwok/Grunfeld 1996). TREC is an annual meeting in which a large real world test collection as a common testbed for IR systems is provided.

Neither Kohonen nor Hopfield networks include hidden units. Consequently, they cannot exploit the advantages of subsymbolic processing like the backpropagation algorithm. Backpropagation is seldomly part of an IR system and it is usually not used for the central task in IR which is the match between query and document, but rather for other tasks as in Crestani/Rijsbergen 1997. Although these systems are based on neural networks, they very much resemble the standard vector space model. In Mothe 1994 partial formal equivalence of both was proved.

3.3.2 The COSIMIR Model

The COSIMIR (COgnitive SIMilarity learning in Information Retrieval) model (Mandl 1998a) intends to use backpropagation for the central problem in IR, to calculate the similarity between query and document as a measure for relevance. In most cases, mathematical functions like the cosine or the inner product are used to calculate the similarity in IR. The decision for the use of one or the other can only be made based on empirical data. Jones/Furnas 1987 discuss the sensitivity of various similarity measures to between-object and within-object differences of term weights. However, their analysis provides little guidance for the developer of an IR system. Additionally, Tversky 1977 reports experiments suggesting that the human similarity judgement lacks formal properties such as symmetry and transitivity which mathematical similarity functions embody.

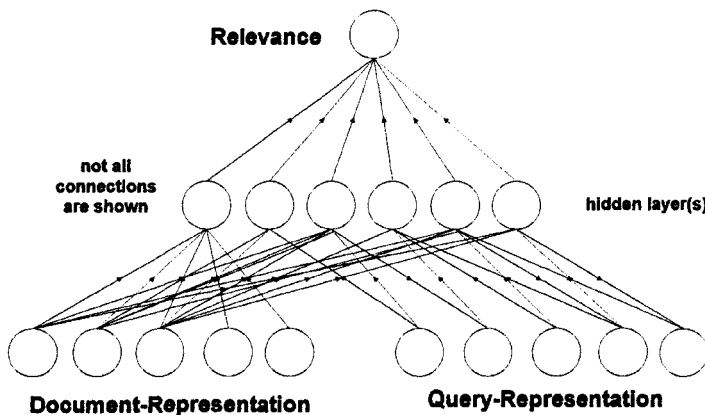


Figure 3: COSIMIR Model

COSIMIR consists of a simple backpropagation neural network which learn the similarity function from users judgements. At the input layer, the query as well as one document representation are presented to the network. The similarity is then calculated in the output unit. Training data for COSIMIR are pairs of documents and queries with relevance (similarity) judgements. The training set also needs to contain cases of little and zero relevance. During the training process, COSIMIR can learn, how the combination of terms in document and query influences relevance. No assumptions on binary independence are necessary, as it is the case in many other IR models.

First experiments with very short vectors from a materials database have shown positive results (see Table). However, experiments with the Cranfield II collection have not produced satisfying results. The number of relevance judgements for the Cranfield II collection is not sufficient for COSIMIR. Further experiments focus on the use of compressed representations thus reducing the number of input nodes and links in the network.

An equivalent model was proposed by de Jong et al. 1996 as cognitive similarity function within a Case Based Reasoning system for industrial processes.

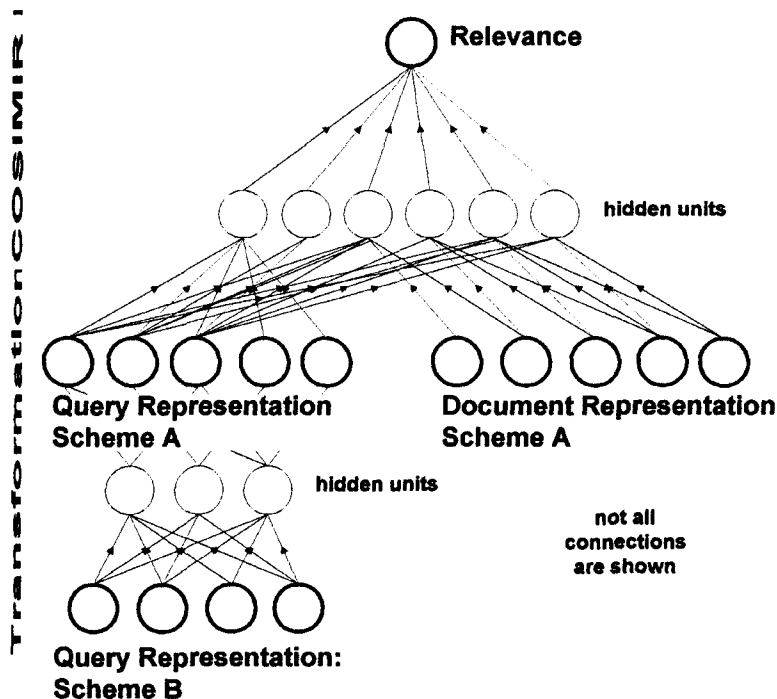


Figure 4: Combination of COSIMIR and a Transformation Network

3.3.3 Adaption of COSIMIR for Heterogeneous IR

COSIMIR can be combined with the transformation network as well as with any other transformation method. Either query or document is transformed and afterwards query and document are compared. Both networks can be combined such that the output of the transformation network will provide part of the input for COSIMIR (fig. 4).

COSIMIR can be adapted to allow queries in heterogeneous datapools, thus requiring only one network. The query and document do not need to be represented within the same scheme as in classical IR systems using mathematical similarity functions. Such a query representation in one scheme can be directly compared to documents indexed differently once sufficient training data has been collected (fig. 5). No explicit transformation step is necessary. Like the other transformation methods, COSIMIR also needs a corpus of comparable documents. However, the relevance judgements can be gradual whereas the other methods require binary decisions on comparability.

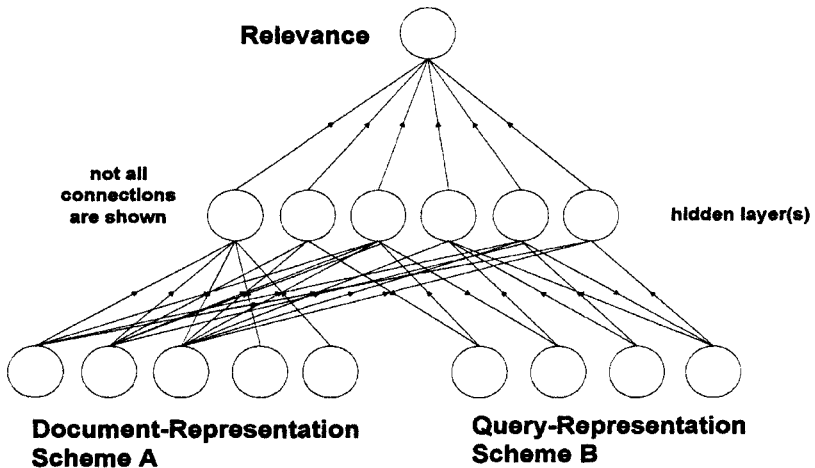


Figure 5: Adapted COSIMIR-Model

This approach was first validated with a small dataset described in Ludwig/Mandl 1997, which represents materials. For each material, two representations were available, namely the physical features and the usage profile. For the users, the usage profile was usually the basis for a similarity judgement. In Ludwig/Mandl 1997, a neural network is discussed which transforms the feature vector into the usage profile. The network reached an overall recognition rate of 85%. As not enough human similarity judgements were available, the similarity was calculated based on the usage profile using a standard similarity function. This heuristic approach seems valid as human judgement is usually based on the usage profile. Three experiments with the same target data could be carried out. First, COSIMIR learned to implement a

mathematical similarity function, receiving the same input as the cosine. Second, a standard COSIMIR network learned to calculate the usage-similarity based on the feature profiles, thus going beyond the pure implementation of a mathematical function. Third, COSIMIR learned to calculate the usage-similarity receiving one feature vector and one usage profile, thus relying on heterogeneous representations. The data set consists of 72 materials of which one third was used for the test set. Input and output vector each have 22 elements.

The evaluation of the results did not consider the absolute value of the similarity as it has little importance in IR. It seemed rather appropriate to compare only the ranked lists, where all materials are ordered according to their similarity to one query material. The average correlation between the in target and result lists were calculated and are shown in the following table.

Table: Results with the Materials Dataset

Target	First Input	Second Input	Output	Correlation
implement mathematical similarity function	usage profile	usage profile	cosine based on usage profiles	79%
standard COSIMIR-model	features	features	cosine based on usage profiles	70%
heterogeneous representations	usage profile	features	cosine based on usage profiles	50%

Although the difficulty of the task reduces the quality of the mapping, the results are satisfying, considering that the correlation between the calculated usage-similarity and the calculated feature-similarity is below 40%. Details on these experiments can be found in Mandl 1998a and Mandl 1998b. Further experiments are nevertheless needed.

4 Conclusions

This article discusses three methods for transformations between heterogeneous representation schemes. Whereas the statistical methods are well tested, the neural networks are still in an experimental stage. However, their capabilities to approximate complex functions and first results with small data sets encourage further investigation.

The construction of a proper knowledge base with double representations for aligned or identical documents is crucial for the success of any vague transformation. In the context of ELVIRA, this is currently a priority. There are not enough user or expert judgements available, therefore aligned corpora need to be constructed heuristically. In one text collection of ELVIRA, many texts contain tables which can be identified. As the tables deal with same topic as the text and may contain terms which are more typical for time series descriptions, text and corresponding table may be used as training pair. In

addition, this text collection uses its own thesaurus, so that relations between its terms and free text can be determined. For a first approach, statistical methods are applied to the ELVIRA data.

References

[Biebricher et al. 1988]

Biebricher, B.; Fuhr, N.; Lustig, G.; Schwantner, M.; Knorz, G.: *The automatic indexing system AIR/PHYS - from research to application*. In: Chiaramella, Yves (ed.) (1988): Proc. of the 11th Int. SIGIR Conf. ACM. New York. p. 333-342.

[Boughanem/Soulé-Dupuy 1998]

Boughanem, M.; Soulé-Dupuy, C.: *Mercure at trec6*. In: Harman, Donna (ed.) (1998): The Sixth Text Retrieval Conference (TREC-6).

[Chen 1995]

Chen, Hsinchun: *Machine Learning for Information Retrieval: Neural Networks, Symbolic Learning, and Genetic Algorithms*. In: Journal of the ASIS. vol. 46(3). p. 194-216.

[Crestani/Rijsbergen 1997]

Crestani, Fabio; Rijsbergen, Cornelis J. van: *A Model for Adaptive Information Retrieval*. In: Journal of Intelligent Information Systems.

[Doszkocs et al. 1990]

Doszkocs, T.E.; Reggia, J.; Lin, X.: *Connectionist Models and Information Retrieval*. In: Annual Review of Information Science and Technology (ARIST), vol. 25. p. 209-260.

[Frei et al. 1996]

Frei, Hans-Peter; Harman, Donna; Schäuble, Peter; Wilkinson, Ross (eds.): *Proc. of the 19th Annual Int. ACM SIGIR Conf. on Information Retrieval*. New York.

[Hull/Grefenstette 1996]

Hull, David; Grefenstette, Gregory: *Querying Across Languages: A Dictionary-Based Approach to Multilingual Information Retrieval*. In: Frei et al. 1996. p. 49-57.

[Jones/Furnas 1987]

Jones, William; Furnas, George: *Pictures of Relevance: A Geometric Analysis of Similarity Measures*. In: Journal of the ASIS. vol. 38(6). p. 420-442.

[Jong et al. 1996)]

Jong, E. de; Keuken, H.; Pol, E. van der; Dekker, E. den; Kerckhoffs, E. J.: *Exergy Analysis of Industrial Processes Using AI Techniques*. In: Computers and Chemical Engineering. vol. 20. p. S1631-S1636.

[Krause 1996]

Krause, Jürgen: *Informationserschließung und -bereitstellung zwischen Deregulation, Kommerzialisierung und weltweiter Vernetzung*. („Schalenmodell“). IZ-Arbeitsbericht 6, IZ Sozialwissenschaften, Bonn.

[Krause et al. 1998]

Krause, Jürgen; Mandl, Thomas; Schaefer, André; Stempfhuber, Maximilian: *Text-Fakten-Integration in Informationssystemen*. In this volume.

[Krause/Womser-Hacker 1997]

Krause, Jürgen; Womser-Hacker, Christa (eds.): *Vages Information Retrieval und graphische Benutzeroberflächen - Beispiel Werkstoffinformation*. Konstanz.

[Kwok/Grunfeld 1996]

Kwok, K. ; Grunfeld, L.: *TREC-4 Ad-Hoc, Routing Retrieval and Filtering Experiments using PIRCS*. In: Harman, Donna (ed.) (1996): *The Fourth Text Retrieval Conference*.

[Ludwig/Mandl 1997]

Ludwig, Michaela; Mandl, Thomas: *Ähnlichkeit von Werkstoffen: Die Anwendung unterschiedlicher Wissensmodellierungstechniken für eine intelligente Komponente von WING*. In: Krause/Womser-Hacker 1997. p. 169-184.

[Mandl 1998a]

Mandl, Thomas: *Das COSIMIR-Modell: Information Retrieval mit Neuronalen Netzen*. Informationszentrum Sozialwissenschaften Bonn, ELVIRA Arbeitsbericht 10. 1998.

[Mandl 1998b]

Mandl, Thomas: *Der Einsatz vager Verfahren für Transformationen*. Informationszentrum Sozialwissenschaften Bonn, ELVIRA Arbeitsbericht 13. 1998.

[Mothe 1994]

Mothe, Josiane: *Search Mechanisms Using a Neural Network Model*. In: *Intelligent Multimedia Information Retrieval Systems and Management. Proceedings of the RIAO '94*. New York. p. 275-294.

[Scheinost et al. 1998]

Scheinost, Ulrich; Haas, Hansjörg; Krause, Jürgen; Lindlbauer, Jürg (eds.): *Marktanalyse und Marktprognose. Das ZVEI Verbandsinformationssystem ELVIRA*. Bonn: Informationszentrum Sozialwissenschaften.

[Sheridan/Ballerini 1996]

Sheridan, Páraic, Ballerini, Jean Paul: *Experiments in Multilingual Information Retrieval using the SPIDER System*. In: Frei et al. 1996. p. 58-65.

[Smolensky 1988]

Smolensky, Paul: *On the Proper Treatment of Connectionism*. In: *Behavioral and Brain Sciences*. vol. 11. 1988. p. 1-74.

[Tversky 1977]

Tversky, Amos: *Features of Similarity*. In: *Psychological Review*. vol. 84(4) July. p. 327.

[Zell 1994]

Zell, Andreas: *Simulation Neuronaler Netze*. Bonn et al.

Section 5

Recent Developments of Information Systems and Tools

(Quality Aspects)

Neuere Entwicklungen auf dem Gebiet temporaler Datenbanken - eine kritische Analyse

Alexander Kaiser

Dr. Alexander Kaiser
Abteilung Angewandte Informatik
Institut für Informationsverarbeitung- und wirtschaft
Wirtschaftsuniversität Wien
Augasse 26, 1090 Wien
EMail: alexander.kaiser@wu-wien.ac.at

Inhalt:

1. Einführung und Problemstellung
2. Temporale Datenmodelle und temporale Datenbanksprachen
3. Ausgewählte temporale Datenmodelle
 - 3.1 Das BCDM und TSQL2
 - 3.2 Das TRM und TSQL
 - 3.3 Das Temporal Relational Data Model und SQL^T
 - 3.4 IXRM und IXSQL
 - 3.5 ATSQL2
 - 3.6 SQL/Temporal, der zeitbezogene Teil des SQL3 Standardentwurfs
4. Bewertung und Vergleich der Modelle
 - 4.1 Bewertung der vorgestellten Modelle
 - 4.2 Ein vergleichender Überblick der vorgestellten Modelle
5. Die Abbildung zeitbezogener Daten im relationalen Modell
 - 5.1 Temporale Integritätsbedingungen
 - 5.2. Attribute mit Zeitbezug
6. Zusammenfassung und Schlußfolgerungen

Zusammenfassung:

Der vorliegende Beitrag skizziert die wichtigsten temporalen Datenmodelle und temporalen Datenbanksprachen, vergleicht sie miteinander und beurteilt sie in Bezug auf Vollständigkeit, Aufwärtskompatibilität zum relationalen Datenmodell und Verfügbarkeit in der Praxis. Diesen Modellen wird ein Ansatz gegenübergestellt, der ohne das relationale Modell zu erweitern eine effiziente Verwaltung zeitbezogener Daten vorstellt. Insbesondere wird in diesem Ansatz auf den Entwurf von Datenbanken mit zeitbezogenen Daten im Detail eingegangen.

1. Einführung und Problemstellung

Temporale Datenbanken sind in den letzten Jahren zu einem bedeutenden Forschungsbereich auf dem Gebiet der Datenbankforschung geworden. Das zunehmende Interesse an temporalen Datenbanken hat mehrere Ursachen. Zum einen ist ein verstärktes Interesse am Thema der Zeit an sich zu beobachten, was unter anderem auch mit dem nahenden Jahrtausendwechsel in Zusammenhang stehen dürfte. Zum anderen ist die Zeit ein wesentlicher Aspekt der realen Welt. Datenbanksysteme sollen letztendlich eine möglichst genaue Abbildung der modellierten Realität darstellen. Spielt die Zeit in der Realität eine bedeutende Rolle, so muß die Zeitdimension daher auch in Datenbanksystemen angemessen berücksichtigt werden. Darüber hinaus hat in den letzten Jahren auch das Konzept des Datawarehouse stark an Bedeutung gewonnen. In Datawarehousesystemen spielt die Zeitdimension eine sehr wichtige Rolle, so daß auch von diesem Bereich her das Forschungsgebiet der temporalen Datenbanken wichtige Impulse erhalten hat.

Ziel des vorliegenden Beitrags ist es, die wichtigsten zeitbezogenen Datenmodelle und Datenbanksprachen kritisch zu beleuchten und sie einem vom Autor der Arbeit entwickelten Ansatz gegenüber zu stellen, der das weit verbreitete relationale Modell als Datenmodell verwendet.

2. Temporale Datenmodelle und temporale Datenbanksprachen

In temporalen Datenbanken wird zwischen der Gültigkeitszeit, das ist die Zeit, die festlegt, in welchem Zeitraum ein Objekt in der modellierten Realität den beschriebenen Zustand aufweist, und der Transaktionszeit, das ist der Zeitpunkt, zu dem ein Datensatz in die Datenbank aufgenommen bzw. geändert oder gelöscht wird, unterschieden. In weiterer Folge der Arbeit beschränken wir uns schwerpunktmäßig auf die Betrachtung der Gültigkeitszeit. Eine weitere wichtige Unterscheidung besteht zwischen der Attributzeitstempelung und der Tupelzeitstempelung. Bei der Attributzeitstempelung wird einem einzelnen Attribut ein Zeitstempel zugeordnet. Dadurch sind die Werte eines Attributs nicht mehr atomar und entsprechen nicht mehr der 1.Normalform. Bei der Tupelzeitstempelung wird einem gesamten Tupel ein Zeitstempel zugeordnet, so daß alle Attributsausprägungen dieses Tupels in diesem Zeitraum in der modellierten Realität gültig sein müssen.

Im folgenden Abschnitt der Arbeit werden einige Modelle kurz vorgestellt. Ziel ist dabei aber keine detaillierte Beschreibung der Modelle, sondern eine Gegenüberstellung der wesentlichen Charakteristika der Modelle und eine Beurteilung der Modelle in Bezug auf Vollständigkeit, kommerzielle Verfügbarkeit und Praktikabilität.

3. Ausgewählte temporale Datenmodelle

Im folgenden werden für jedes der ausgewählten Datenmodelle die jeweils typischen Charakteristika näher beschrieben, also vor allem diejenigen Aspekte, in denen sich das jeweilige Datenmodell von den übrigen Datenmodellen unterscheidet. Für eine vollständige Beschreibung der Datenmodelle und Datenbanksprachen sei auf die jeweils angegebene Literatur verwiesen. Die Auswahl der Datenmodelle erfolgte nach den Aspekten Aktualität (ATSQL2, SQL/Temporal), Repräsentanten der Tupelzeitstempelung (TSQL2, TSQL, IXSQL, ATSQL2, SQL/Temporal) und Repräsentanten der Attributzeitstempelung (SQL^T).

3.1 Das BCDM und TSQL2

Das BCDM (Bitemporal Conceptual Data Model) und die darauf basierende Sprache TSQL2 wurde von einer größeren Gruppe von Forschern entwickelt, die sich mit temporalen Daten beschäftigen (Snodgrass et al. 95). Das BCDM ist ein bitemporales Datenmodell mit Tupelzeitstempelung, d.h. es unterstützt sowohl die Gültigkeitszeit als auch die Transaktionszeit. Das BCDM sieht ein implizites Zeitstempelattribut vor und steht damit im Gegensatz zu den meisten anderen tupelorientierten Modellen, bei denen die Zeitstempelattribute explizite Attribute sind. Durch die implizite Tupelzeitstempelung wird auch angenommen, daß alle Attribute eines Tupels zeitabhängig sind und sich (im Idealfall) auch gleichzeitig ändern sollten. TSQL2 ist aufwärtskompatibel zu SQL92.

Neben den in SQL92 vorhandenen Datentypen sieht TSQL2 auch den Datentyp PERIOD vor, mit dem eine Periode von einem Beginnzeitpunkt bis zu einem Endzeitpunkt abgebildet werden kann. Der Datentyp PERIOD entspricht im wesentlichen dem Datentyp INTERVAL OF DATE in IXSQL. Der Datentyp PERIOD findet sich auch in SQL/Temporal des SQL3 Standardentwurfs wieder. Für den neuen Datentyp PERIOD werden spezielle Vergleichsoperatoren eingeführt, die eine Untermenge der 13 Zeitoperatoren von Allen (Allen 83) bilden (Myrach 97). Durch den Zusatz STATE bzw. EVENT bei der CREATE TABLE Anweisung in TSQL2 kann man explizit zwischen Zustands- bzw. Ereignis-Tabellen wählen. In Kombination mit der Angabe, ob es sich um eine Schnappschußtabelle (also eine konventionelle Tabelle ohne zeitbezogene Daten), eine bitemporale Tabelle, eine Gültigkeitszeit-Tabelle oder eine Transaktionszeit-Tabelle handelt, sieht TSQL2 sechs verschiedene Typen von Tabellen vor.

Ein wesentliches Charakteristikum des BCDM und TSQL2 ist das automatische Zusammenfassen mehrerer aufeinander folgender Versionen mit identischen Attributwerten, was sowohl in der Literatur als auch in der Praxis nicht unumstritten ist. Dieser Vorgang wird als *coalescing* bezeichnet. Beim coalescing ist die Gefahr eines Informationsverlustes gegeben, da bei der Zusammenfassung zweier aufeinander folgender Versionen zu einer Version, in die Historie eingegriffen wird. Auf der anderen Seite ist die Bearbeitung von Relationen, in denen Versionen existieren, die zusammengefaßt werden können, bei Datenbankabfragen wesentlich schwieriger und aufwendiger. Das coalescing sollte daher bestenfalls explizit unterstützt werden.

Auf Integritätsbedingungen, die gerade bei zeitbezogenen Daten ein Problem darstellen können, wird in TSQL2 überhaupt nicht eingegangen, so daß die Ausführung der Datenmodifikationsoperationen INSERT, UPDATE und DELETE in TSQL2 auch zu logisch widersinnigen bzw. falschen Ergebnissen führen könnten.

Eine gute, komprimierte und auch kritische Beschreibung von TSQL2 findet sich in (Myrach 97).

3.2 Das TRM und TSQL

Bereits etwas älter ist das TRM (Temporal Relational Model) und die darauf basierende Sprache TSQL, entwickelt von Navathe und Ahmed (Navathe and Ahmed 89), (Navathe and Ahmed 93). Das TRM ist eine Weiterentwicklung des relationalen Modells, unterstützt nur die Gültigkeitszeit, und bildet die Zeitdimension mittels Tupelzeitstempelung ab. Insbesondere wird im TRM das Konzept der temporalen Abhängigkeit zwischen zwei zeitabhängigen Attributen vorgestellt. Eine temporale Abhängigkeit ist dann gegeben, wenn sich in einer Relation zwei oder mehrere zeitabhängige Attribute befinden, deren Werte sich nicht zum selben Zeitpunkt ändern. Beim Entwurf temporaler Datenbanken mit dem TRM muß darauf geachtet werden, daß temporale Abhängigkeiten vermieden werden, da es ansonsten zu Update- und Retrievalanomalien kommen kann. Ändern sich zeitabhängige Attribute immer zum selben Zeitpunkt, spricht man von synchronen Attributen. In einer zeitbezogenen Relation können im TRM nur synchrone Attribute enthalten sein. In der folgenden Relation existiert eine temporale Abhängigkeit zwischen den Attributen Funktion und Gehalt, da sich zum Zeitpunkt 1. September 1996 zwar der Gehalt des Mitarbeiters 1 ändert, nicht aber seine Funktion.

<u>P#</u>	<u>Funktion</u>	<u>Gehalt</u>	<u>Beginn</u>	<u>Ende</u>
1	Assistent	20.000	1995-SEP-01	1996-AUG-31
1	Assistent	22.000	1996-SEP-01	1998-DEC-31

Die Relation müßte auf zwei Relationen mit den Attributen P#, Funktion, Beginn, Ende bzw. P#, Gehalt, Beginn, Ende aufgeteilt werden. Die temporale Abhängigkeit zwischen den Attributen Funktion und Gehalt ist deshalb problematisch, da dadurch die Gültigkeitsspanne (life span) des Attributs Funktion auf zwei oder mehrere Tupel aufgeteilt wird und somit einfache Abfragen wie z.B. "Wie lange war ein Mitarbeiter Assistent?" nur mit relativ aufwendigen Abfragen beantwortet werden können. Auch bei der Erfassung neuer Versionen kann es bei Relationen mit temporal abhängigen Attributen zu Problemen kommen. Soll eine neue Version eingefügt werden, da sich z.B. der Gehalt von 22.000 auf 23.000 ändert, muß vor der INSERT Anweisung eine SELECT Anweisung durchgeführt werden, um den aktuellen Wert des Attributs Funktion zu erhalten.

Navathe und Ahmed definieren eine eigene time normal form (TNF), indem sie festlegen, daß eine zeitbezogene Relation dann in TNF ist, wenn sie in BCNF ist und keine temporalen Abhängigkeiten zwischen ihren Attributen existiert. Die Sprache TSQL ist eine Erweiterung von Standard-SQL und erweitert SQL um einige zeitbezogene Konstrukte. Obwohl TSQL ein Gültigkeitsintervall mit

zwei Zeitstempelattributen abbildet, können in TSQL mit dem Konstrukt `Relationenname.INTERVAL` das Gültigkeitsintervall direkt angesprochen und die Zeitoperatoren von Allen (Allen 83) darauf angewendet werden. Mit der `BREAK` Klausel in einer `SELECT` Anweisung können in TSQL auch Lücken in der Historie eines Objekts angesprochen werden. Eine Lücke ist dann gegeben, wenn die Gültigkeitsintervalle der Versionen eines Objekts nicht direkt aneinander anschließen, sondern dazwischen Zeitpunkte existieren, zu denen das Objekt keinen definierten Zustand aufweist. Auch das Gruppieren von zeitbezogenen Daten ist mittels der `MOVING WINDOW` Klausel möglich.

3.3 Das Temporal Relational Data Model und SQL^T

Die Sprache SQL^T (Tansel 93), (Tansel 97a), (Tansel 97b) basiert auf einem Datenmodell mit Attributzeitstempelung. Das heißt insbesondere, daß Relationen verwendet werden, die nicht in 1.Normalform sind (NFNF), da die Werte ihrer zeitabhängigen Attribute nicht atomar sind. Das Datenmodell ist ein Gültigkeitszeitmodell. Das der Sprache SQL^T zugrundeliegende temporale relationale Datenmodell unterstützt "nested relations", das sind Relationen, deren Attribute wieder aus Relationen oder Mengen bestehen können. Das hier behandelte Datenmodell sieht allerdings nur eine Schachtelungstiefe von 1 vor, d.h. daß Attribute entweder atomar sein können, oder aber Mengen sein können, jedoch keine Mengen von Mengen. Jeder Attributwert zeitabhängiger Attribute wird als $\langle \text{Zeit}, \text{Wert} \rangle$ -Paar aufgezeichnet, wobei *Zeit* ein Zeitpunkt sein kann, zu dem der Attributwert belegt wurde, oder ein Intervall während dessen der Attributwert in der modellierten Realität gültig ist.

Die Sprache SQL^T wird nur in (Tansel 93) beschrieben. In seinen späteren Arbeiten zu diesem Thema ((Tansel 97a), (Tansel 97b)) geht der Autor auf diese Sprache nicht mehr ein, obwohl im Schlußteil von (Tansel 93) einige Erweiterungs- und Verbesserungsvorschläge für SQL^T gemacht werden. In (Tansel 93) wird lediglich der Abfrageaspekt der Sprache genau beschrieben. Durch das Voranstellen des Zeichens # bei einem Attribut in der `SELECT` Klausel einer `SELECT` Anweisung wird die gesamte Historie dieses Attributs ausgegeben. Wenn nicht die gesamte Historie eines zeitbezogenen Attributs ausgegeben werden soll, sondern nur ein bestimmter Zeitbereich, kann in der `SELECT` Klausel das Attribut mit einem Zeitzusatz versehen werden. Sei *A* ein zeitabhängiges Attribut und sei *t* eine Menge von Zeitpunkten, so kann mit der Anweisung `SELECT #A[t] FROM Tabelle` der Teil der Historie von *A* ausgegeben werden, der sich innerhalb des angegebenen Zeitraums *t* befindet. Die `COLLAPSE` Klausel dient dazu um mehrere $\langle \text{Zeit}, \text{Wert} \rangle$ -Paare mit aneinander angrenzenden oder überlappenden Zeitbereichen zu einem $\langle \text{Zeit}, \text{Wert} \rangle$ -Paar zusammenzufassen und entspricht damit einem coalescing auf Attributebene.

In seinen späteren Arbeiten (Tansel 97a), (Tansel 97b) definiert Tansel einen temporalen Relationenkalkül und eine temporale relationale Algebra, ohne diese jedoch in eine konkrete temporale Datenbanksprache syntaktisch umzusetzen.

3.4 IXRM und IXSQL

Das Interval extended relational model (IXRM) und die darauf basierende Sprache IXSQL (Lorentzos and Mitsopoulos 97), (Lorentzos 93) wurde von den Griechen Lorentzos und Mitsopoulos entwickelt. IXRM ist ein Gültigkeitszeitmodell, unterscheidet sich jedoch von den meisten anderen Modellen und Sprachen, da im Mittelpunkt dieses Modells die explizite Unterstützung eines Datentyps *Interval* steht. Basierend auf diesem Intervalldatentyp wurden neue Prädikate und Funktionen entwickelt und auch die relationale Algebra intervallmäßig erweitert. IXSQL ist eine aufwärtskompatible Sprache zu SQL92, d.h. daß mit IXSQL auch nicht-zeitbezogene Daten verarbeitet werden können und alle Elemente des SQL92-Standards unterstützt werden. Die wichtigsten Erweiterungen von IXSQL gegenüber SQL92 sind neben dem Intervall-Datentyp die Funktionen *FOLD* bzw. *UNFOLD* und *NORMALIZE*. Mit der Funktion *UNFOLD* ist es möglich, eine Relation von ihrer Intervalldarstellung auf die Darstellung einzelner Zeitpunkte "herunterzubrechen". Wendet man auf folgende Relation die Funktion *UNFOLD* an:

P#	Funktion	Zeit
1	Assistent	[d32,d35)
1	Dozent	[d35,d38)

dann erhält man als Ergebnis die nachstehende Relation:

P#	Funktion	Zeit
1	Assistent	d32
1	Assistent	d33
1	Assistent	d34
1	Dozent	d35
1	Dozent	d36
1	Dozent	d37

UNFOLD ermöglicht somit, ein "umgekehrtes coalescing" durchzuführen. *FOLD* ist die inverse Funktion zu *UNFOLD*. Die Funktion *NORMALIZE* führt zuerst ein *FOLD* und dann auf die so erzeugte Relation ein *UNFOLD* durch.

IXSQL unterstützt alle Datentypen *D* aus SQL92 und zusätzlich für jeden Datentyp *D* auch den korrespondierenden Intervalltyp *I(D)*. Das bedeutet, daß Intervalle auch von Zahlen und Zeichendatentypen und nicht nur von zeitbezogenen Datentypen gebildet werden können. Intervalle sind in IXSQL immer halboffene Intervalle.

Bei den Integritätsbedingungen wird in den Arbeiten zu IXSQL nur auf die Entitätsintegrität Bezug genommen. Der *PRIMARY KEY* einer zeitbezogenen Relation wird gegenüber dem *PRIMARY KEY* der korrespondierenden konventionellen Relation um das Zeitattribut erweitert. Dadurch ist eine effektive Überprüfung der Entitätsintegrität gewährleistet, da mit dem *INTERVAL* Datentyp auch die Überlappung von Zeitintervallen überprüft werden kann.

3.5 ATSQL2

ATSQL2 (Steiner 98) ist eine der neuesten Entwicklungen auf dem Gebiet temporaler Datenbanksprachen. ATSQL2 basiert auf dem relationalen Datenmodell und hat sich als Hauptziel eine konsequente Aufwärtskompatibilität zu SQL92 gesetzt. Im Unterschied zu den meisten anderen Sprachentwürfen existiert für ATSQL2 mit TimeDB 2.0 (TimeDB 98) auch ein bitemporales relationales DBMS, das Anbindungen an kommerziell verfügbare Datenbankmanagementsysteme wie Oracle oder Sybase vorsieht. Die Grundidee liegt dabei darin, daß zwar mit ATSQL2 eine zu SQL92 aufwärtskompatible Sprache definiert wurde, daß jedoch jede ATSQL2-Anweisung in eine gültige SQL92-Anweisung umgewandelt wird. Damit dient ATSQL2 quasi als äußere Schicht, auf der die Benutzer zeitbezogene Daten effizient bearbeiten und abfragen können. Diese Abfragen werden dann in teilweise sehr komplexe SQL92-Befehle umgewandelt und an ein konventionelles DBMS weitergegeben, ohne daß der Anwender diesen "Umwandlungsschritt" bemerkt. Zwei wesentliche Anforderungen an die Aufwärtskompatibilität von ATSQL2 sind, daß jede korrekte SQL92-Anweisung auch in ATSQL2 gültig ist und daß die Syntax von ATSQL2 so ähnlich wie möglich der Syntax von SQL92 ist. Diesen Anforderungen entsprechend werden in ATSQL2 durch das Voranstellen der Schlüsselwörter `VALID` oder `TRANSACTION` oder `VALID AND TRANSACTION` vor einer `SELECT` Anweisung aus konventionellen Abfragen zeitbezogene Abfragen. Die Sprache ATSQL2 unterstützt Gültigkeitsintervalle, wobei diese Intervalle bei der Umwandlung in SQL92 in zwei versteckte Zeitstempelattribute `Begin` und `End` umgewandelt werden. TimeDB hat eine temporale relationale Algebra implementiert, was bedeutet, daß die Standardoperationen aus der relationalen Algebra unter Berücksichtigung der Gültigkeitszeit erweitert und anschließend wieder in SQL92-Anweisungen umgewandelt werden. Auch eine bitemporale relationale Algebra ist vorgesehen, die nach demselben Schema funktioniert. Zusätzlich zu den beiden versteckten Gültigkeitszeitstempelattributen werden dann noch zwei versteckte Transaktionszeitstempelattribute in SQL92 angelegt. Ein wichtiges Konzept von ATSQL2 sind abgeleitete Tabellen. Dabei geht es darum, daß in der `FROM` Klausel einer `SELECT` Anweisung anstelle eines Tabellennamens wieder eine `SELECT` Anweisung stehen kann, deren Ergebnis dann wie eine Tabelle behandelt wird. Dieses Konzept ist auch im Full Level von SQL92 vorgesehen. Auf Integritätsbedingungen wird in den Arbeiten zu ATSQL2 nur kurz eingegangen. Durch das Voranstellen des Schlüsselwortes `VALID` werden aus konventionellen Integritätsbedingungen temporale Integritätsbedingungen. Details dazu sind in den Arbeiten allerdings nicht vorhanden.

3.6 SQL/Temporal, der zeitbezogene Teil des SQL3 Standardentwurfs

Seit einigen Jahren wird an der Weiterentwicklung des derzeit gültigen SQL-Standards gearbeitet. Der Entwurf für diese Weiterentwicklung wird als SQL3 (ISO/IEC 97) bezeichnet und enthält mit SQL/Temporal auch einen Teil, der sich mit zeitbezogenen Daten beschäftigt. Das zugrundeliegende Datenmodell bleibt freilich das relationale Datenmodell, so daß bei SQL/Temporal in erster Linie Erweiterungen in Bezug auf die benutzerdefinierte Zeit enthalten sind. Wesentlichste Neuerung ist die Einführung eines neuen Datentyps `PERIOD`. `PERIOD` entspricht dem Datentyp `INTERVAL OF DATE` in IXSQL. SQL/Temporal

besitzt damit mit DATE, TIME, TIMESTAMP, INTERVAL und PERIOD Datentypen für die Abbildung von Zeitpunkten, von Perioden und von Intervalldauern. Mit CONTAINS, PRECEDES, SUCCEEDS, OVERLAPS und MEETS werden auch alle für den Datentyp PERIOD relevanten Operationen definiert. Darüber hinaus wird in SQL/Temporal auch eine NORMALIZE Funktion definiert. Die Aufgabe dieser Funktion ist dieselbe wie in IXSQL.

4. Bewertung und Vergleich der Modelle

In diesem Abschnitt der Arbeit werden die Modelle vorerst einzeln beurteilt und daran anschließend basierend auf sieben ausgewählten Kriterien miteinander verglichen.

4.1 Bewertung der vorgestellten Modelle

BCDM und TSQL2: Beim Sprachentwurf von TSQL2 werden eine große Anzahl impliziter Annahmen getroffen, deren Ausnutzung die Formulierung kompakter Anfragen erlaubt (Myrach 97). Problematisch erscheint an diesem Ansatz, daß der Aspekt des Datenbankentwurfs völlig vernachlässigt wird. Das geht soweit, daß in TSQL2 keine Konstrukte für Integritätsbedingungen, wie PRIMARY KEY oder FOREIGN KEY, explizit vorgesehen sind. Das automatische coalescing kann ebenso wie beim TRM und TSQL zu massiven Informationsverlusten führen.

TRM und TSQL: Das Konzept der temporalen Abhängigkeit hilft beim effizienten Entwurf einer temporalen Datenbank. Problematisch ist jedoch, daß mit diesem Modell zeitabhängige Attribute, deren Werte sich über mehrere Perioden eines Beobachtungszeitraums nicht ändern, nicht explizit abgebildet werden können. Das führt zu einem Informationsverlust, nämlich einem Verlust der Information, daß sich der Wert eines Attributs eben nicht geändert hat. Auf die referentielle Integrität wird im TRM nicht eingegangen. Bei der Beschreibung von TSQL wird ein Hauptgewicht auf die erweiterte SELECT Anweisung gelegt. Die Aspekte der Datendefinitionssprache und der Datenmanipulationsoperationen werden hingegen völlig vernachlässigt.

SQL^T: Die Attributzeitstempelung hat gegenüber der Tupelzeitstempelung unter anderem den Vorteil, daß die gesamte Historie eines zeitabhängigen Attributs direkt beim Attribut abgespeichert ist. Deshalb kann bei Modellen mit Attributzeitstempelung das "Coalescing-Problem" (vgl. dazu Abschnitt 3.1 dieser Arbeit) nie auftreten. Auf das Design zeitbezogener Relationen mit SQL^T und die Datenmanipulation mit dieser Sprache wird in den Arbeiten von Tansel überhaupt nicht eingegangen. Daher ist es auch sehr schwierig, die Behauptung des Autors zu überprüfen, daß SQL^T zu Standard-SQL aufwärtskompatibel ist. Jedenfalls bleibt die Tatsache bestehen, daß das Datenmodell, auf dem SQL^T basiert, zum relationalen Datenmodell, auf dem die meisten weit verbreiteten DBMS basieren, nicht aufwärtskompatibel ist. Dieses Manko macht eine baldige Verfügbarkeit von Sprachen mit Attributzeitstempelung in kommerziell verfügbaren DBMS schwer möglich. Ein alternativer Ansatz ist m.E. darin zu se-

hen, Datenmodelle mit Attributzeitstempelung im allgemeinen und das Datenmodell von SQL^T im besonderen in objektrelationalen Datenmodellen abzubilden. Mit Oracle 8.0 unterstützt ein weit verbreitetes kommerzielles DBMS ein objektrelationales Datenmodell.

IXRM und IXSQL: Der wesentliche Vorteil von IXSQL liegt in der expliziten Unterstützung von Intervallen und der Definition von intervallspezifischen Funktionen und Prädikaten. Gültigkeitsintervalle können damit mit einem Attribut abgebildet werden und brauchen nicht mehr mit zwei Zeitstempelattributen (Beginn und Ende) quasi simuliert werden. Das hat unter anderem auch den Vorteil, daß das DBMS automatisch die ordnungsgemäße Verwendung von Intervallen überprüft. Beispielsweise ist es nicht möglich, ein Intervall anzugeben, bei dem die untere Grenze größer als die obere Grenze ist, was bei der Darstellung von Intervallen mit zwei Zeitstempelattributen erst durch entsprechende Integritätsbedingungen explizit ausgeschlossen werden muß. Weiters ist es in IXSQL auch möglich, mehrere Gültigkeitszeitintervalle anzulegen. Auf der Ebene der Integritätsbedingungen wird nur auf die Entitätsintegrität näher eingegangen. Mit der expliziten Erweiterung des PRIMARY KEY um das Intervall-Zeitattribut ist eine effektive Überwachung der Entitätsintegrität möglich, da auch die Überlappung von Zeitintervallen überprüft wird. Auf die referentielle Integrität wird in den Arbeiten von Lorentzos und Mitsopoulos leider nicht eingegangen. Positiv zu vermerken ist noch, daß in IXSQL ein coalescing, das dort als NORMALIZE bezeichnet wird, zwar möglich ist, vom Datenbankdesigner aber explizit angegeben werden muß. D.h., es gibt kein implizites Zusammenfassen von Datensätzen wie in TSQL2.

ATSQL2: Der Vorteil dieses Ansatzes liegt vor allem darin, daß diese Sprache mit kommerziell verfügbaren DBMS kombinierbar und vollständig aufwärtskompatibel zu SQL92 ist. Damit ist ein schrittweiser Übergang von einer konventionellen Datenbank auf eine temporale Datenbank möglich, ohne daß "Altsysteme" neu programmiert werden müssen. Da die Sprache kein eigenes Datenmodell hat, sondern auf dem relationalen Modell basiert, werden alle ATSQL2-Anweisungen in SQL92-Anweisungen umgewandelt. Die dadurch entstehenden SQL92-Anweisungen sind oftmals sehr komplex, was negative Auswirkungen auf die Performance haben kann. Angaben dazu liegen in den aktuellen Arbeiten jedoch noch nicht vor.

SQL Temporal: In SQL/Temporal sind Konzepte und Ideen von TSQL2, ATSQL2 und IXSQL eingeflossen. SQL/Temporal kann damit als ein guter Kompromiß der wichtigsten Sprachentwürfe verstanden werden, da sich in dieser Sprache eine große Anzahl von Forschern auf dem Gebiet temporaler Datenbanken wiederfindet. Darüber hinaus sollte SQL/Temporal als Teil des neuen SQL3-Standards nach dessen Verabschiedung auch für die Hersteller allgemein verbindlich werden. Ähnlich wie bei ATSQL2 kann auch SQL/Temporal als "zeitbezogene Schicht" eines DBMS verstanden werden, da das zugrundeliegende Datenmodell mit dem relationalen Modell nach wie vor kein temporales Datenmodell ist.

4.2 Ein vergleichender Überblick der vorgestellten Modelle

In der folgenden Übersicht sollen die vorgestellten Datenmodelle und Datenbanksprachen nochmals miteinander verglichen werden. Für diesen Vergleich werden sieben Vergleichskriterien (V1-V7) herangezogen. Die Beschreibung einiger Datenmodelle bzw. Datenbanksprachen in der Literatur ist nicht vollständig, so daß nicht alle Vergleichskriterien eindeutig bestimmt werden können.

V1: Welche Zeitdimension wird unterstützt?

V2: Wird die Entitätsintegrität angemessen unterstützt?

V3: Wird die referentielle Integrität angemessen unterstützt?

V4: Wird ein coalescing automatisch durchgeführt?

V5: Ist das Datenmodell zum relationalen Modell aufwärts kompatibel?

V6: Ist die Datenbanksprache zu SQL92 (syntaktisch) aufwärts kompatibel?

V7: Gibt es einen (kommerziell) verfügbaren Prototyp der Sprache?

	<u>TSQL2</u>	<u>TSQL</u>	<u>SQL^T</u>	<u>IXSQL</u>	<u>ATSQL2</u>	<u>SQL/Temporal</u>
V1	bi	Gültigk.	Gültigk.	Gültigk.	bi	Gültigk.
V2	nein	ja ¹	k.A. ²	ja	ja? ³	k.A. ²
V3	nein	nein	k.A. ²	nein	ja? ³	k.A. ²
V4	ja	ja	nein	nein	nein	nein
V5	nein	ja	nein	ja	ja	ja
V6	ja	ja	ja	ja	ja	ja
V7	nein	nein	nein	nein	ja	nein

Ein Mangel ist fast allen vorgestellten Modellen und auch allen anderen in der Literatur behandelten temporalen Datenmodellen gemeinsam: die Modelle und die darauf basierenden Datenbanksprachen sind in keinen kommerziell verfügbaren DBMS implementiert. Auch mittelfristig ist nicht zu erwarten, daß ein temporales Datenbankmanagementsystem verfügbar sein wird. Aus diesem Grund erscheint es sinnvoll, sich verstärkt mit der Frage auseinanderzusetzen, wie die Zeitdimension in bereits bestehende und bewährte Modelle integriert werden kann.

In (Kaiser 96) wurden dazu bereits einige Gedanken dargestellt. Im folgenden Abschnitt wird auf neue Erkenntnisse und Weiterentwicklungen mit diesem Ansatz näher eingegangen. Insbesondere wurde in den jüngsten Arbeiten an diesem Modell versucht, auch Aspekte von Modellen mit Attributzeitstempelung zu berücksichtigen, dabei aber trotzdem im Rahmen des konventionellen relatio-

¹ Allerdings nur im Datenmodell und nicht in der Datenbanksprache

² k.A. = in der Literatur wurden dazu keine entsprechenden Angaben gefunden.

³ In den Arbeiten zu ATSQL2 wird zwar von entsprechenden Integritätsbedingungen gesprochen. Eine genauere Beschreibung derselben erfolgt aber nicht, so daß hier keine eindeutige Antwort gegeben werden kann.

nalen Modells zu bleiben, um eine Implementierung in weit verbreiteten Datenbankmanagementsystemen zu ermöglichen.

5. Die Abbildung zeitbezogener Daten im relationalen Modell

Die Analyse und Bewertung der vorgestellten Modelle hat gezeigt, daß die Modelle vor allem in der Phase des Entwurfs temporaler Datenbanken keine angemessene Unterstützung bieten. Kein einziges Modell behandelt alle Integritätsbedingungen in temporalen Datenbanken zufriedenstellend. Auch in der aktuellsten Literatur wird auf diesen so wichtigen Aspekt kaum eingegangen (eine Ausnahme dabei bildet (Bergamaschi and Sartori 98)). Im folgenden wollen wir uns daher auf den Aspekt des Entwurfs temporaler relationaler Datenbanken konzentrieren.

Ein wichtiger Schwerpunkt bei der Abbildung zeitbezogener Daten im relationalen Modell muß in der sauberen Definition adäquater Integritätsbedingungen liegen. Dazu werden mit dem PRIMARY KEY TEMPORAL (Kaiser 98b) und dem FOREIGN KEY TEMPORAL (Kaiser 98a) zwei spezifisch zeitbezogene Integritätsbedingungen eingeführt. Diese beiden Konstrukte stellen zwar Erweiterungen zu SQL92 dar, können jedoch auch mit dem im SQL92-Standard definierten Sprachumfang formuliert werden. Das Konzept der temporalen Abhängigkeit von Attributen wird aufgenommen und erweitert. Dabei wird jedoch zwischen verschiedenen Kategorien von zeitabhängigen Attributen unterschieden. Durch diese Unterscheidung kann wesentlich besser auf den Kontext der abgebildeten modellierten Realität Rücksicht genommen werden. Obwohl das Modell eine Tupelzeitstempelung verfolgt, können durch die Unterscheidung in unterschiedliche Attributkategorien auch einige Aspekte der Attributzeitstempelung einfließen. Im folgenden werden die Ecksteine dieses Modells kurz vorgestellt.

5.1 Temporale Integritätsbedingungen

Die PRIMARY KEY TEMPORAL Klausel ist eine Erweiterung von SQL92 und stellt die temporale Entitätsintegrität sicher. Die temporale Entitätsintegrität legt fest, daß zu jedem Zeitpunkt t jedes Objekt nur eine gültige Version besitzen darf. Wie aus der folgenden Definition ersichtlich ist, kann die PRIMARY KEY TEMPORAL Klausel mittels einer expliziten Integritätsbedingung auch in SQL92 aufgelöst werden.

Seien die c_i die Komponenten des PRIMARY KEY TEMPORAL.

Die PRIMARY KEY TEMPORAL-Bedingung ist genau dann verletzt, wenn folgende Suchbedingung wahr ist:

```
EXISTS (SELECT * FROM B a, B b
        WHERE NOT (UNIQUE (SELECT a.c1, ..., a.cn FROM a)
                   AND (a.c1, ..., a.cn) IS NOT NULL
                   OR ((a.c1, ..., a.cn) = (b.c1, ..., b.cn) AND
                      (a.TB <> b.TB) AND
                      (A.TB = b.TE OR a.TE = b.TB OR
                      (a.TB, a.TE) OVERLAPS (b.TB, b.TE))))
```

Die FOREIGN KEY TEMPORAL Klausel ist eine Erweiterung zu SQL92 und stellt die temporale referentielle Integrität sicher. Die temporale referentielle Integrität

legt fest, daß eine Version eines Objekts sich nur dann auf eine referenzierte Tabelle beziehen kann, wenn die Gültigkeitsperiode des Tupels in der referenzierten Tabelle die Gültigkeitsperiode des Tupels in der referenzierenden Tabelle beinhaltet. Auch die FOREIGN KEY TEMPORAL Klausel kann mittels einer expliziten Integritätsbedingung in SQL92 aufgelöst werden.

Seien R_2 eine Tabelle und FK der Fremdschlüssel der Tabelle und T_B, T_E die Zeitattribute dieser Tabelle. R_2 ist die referenzierende Tabelle. Seien R_1 eine Tabelle und PK, T_B die Attribute des PRIMARY KEY TEMPORAL dieser Tabelle und T_B, T_E die Zeitattribute von R_1 . R_1 ist die referenzierte Tabelle.

Die FOREIGN KEY TEMPORAL-Bedingung ist erfüllt, wenn die folgende Suchbedingung wahr ist:

```
EXISTS (SELECT * FROM R2, R1
        WHERE FK=PK
        AND R2.TB >= R1.TB
        AND R2.TE <= R1.TE);
```

Mit diesen beiden Konstrukten ist eine effiziente und saubere Überwachung der Entitätsintegrität und der referentiellen Integrität von Tabellen mit zeitbezogenen Daten gewährleistet.

5.2. Attribute mit Zeitbezug

Bereits die Beschreibung und Bewertung der Datenmodelle in Abschnitt 3 und 4 dieser Arbeit hat gezeigt, daß die Attributzeitstempelung einige Vorteile gegenüber der Tupelzeitstempelung bietet. Insbesondere ist dabei die direkte Verbindung der Historie mit einem Attribut hervorzuheben. Ziel unserer laufenden Forschungsarbeiten ist es daher, basierend auf dem traditionellen relationalen Modell die Attribute einer zeitbezogenen Relation mit zusätzlichen semantischen Informationen zu versehen. Unter einer zeitbezogenen Relation wird dabei eine konventionelle Relation verstanden, die um zwei Zeitstempelattribute (Beginn und Ende) erweitert wird.

In zeitbezogenen Relationen kann zwischen fünf unterschiedlichen Attributarten unterschieden werden:

- Zeitstempelattribute
- zeitunabhängige Attribute
- zeitabhängige Attribute im weiteren Sinn
- zeitabhängige Attribute im engeren Sinn
- zeitabhängige zyklische Attribute

Diese fünf unterschiedlichen Attributarten übernehmen alle Eigenschaften eines Attributs aus dem relationalen Modell, erhalten aber darüber hinaus zusätzliche semantische Informationen. In weiterer Folge werden die je eigenen Charakteristika dieser Attributarten beschrieben.

Klasse 0: Zeitstempelattribute: Die beiden Zeitstempelattribute T_B und T_E sind auf der Domäne 'Zeit' definiert und legen den Beginn und das Ende einer Gültigkeitsperiode fest.

Klasse 1: zeitunabhängige Attribute: Die Werte von zeitunabhängigen Attributen bleiben im Zeitablauf einer zeitbezogenen Relation konstant. Ein typisches Beispiel für ein zeitunabhängiges Attribut wäre etwa das Geburtsdatum eines Mitarbeiters.

Klasse 2: zeitabhängige Attribute im weiteren Sinn: Die Werte von zeitabhängigen Attributen im weiteren Sinn können sich im Zeitablauf einer zeitbezogenen Relation ändern. Die Wertänderungen dieser Attribute sind jedoch im Kontext der modellierten Realität nicht relevant. Das bedeutet, daß die Historie solcher Attribute nicht nachvollziehbar sein muß und daß immer nur der momentan gültige und aktuelle Wert dieses Attributs relevant ist. Ein Beispiel für ein zeitabhängiges Attribut im weiteren Sinn wäre etwa der Name eines Mitarbeiters, natürlich aber nur dann, wenn frühere Namen eines Mitarbeiters in der modellierten Realität nicht relevant sind.

Klasse 3: zeitabhängige Attribute im engeren Sinn: Die Werte von zeitabhängigen Attributen im engeren Sinn können sich im Zeitablauf einer zeitbezogenen Relation ändern. Die Wertänderungen dieser Attribute sind relevant, d.h. die Historie dieser Attribute muß nachvollziehbar sein. Wird für ein Objekt eine neue Version angelegt, so muß sich der Wert eines zeitabhängigen Attributs im engeren Sinn gegenüber dem Wert in der Vorversion jedenfalls ändern. Ein Beispiel für ein zeitabhängiges Attribut im engeren Sinn wäre etwa die Funktion eines Mitarbeiters.

Klasse 4: zeitabhängige zyklische Attribute: Die Werte von zeitabhängigen zyklischen Attributen können sich im Zeitablauf einer zeitbezogenen Relation ändern. Die Wertänderungen dieser Attribute sind relevant, d.h. die Historie dieser Attribute muß nachvollziehbar sein. Wird für ein Objekt eine neue Version angelegt, so muß sich der Wert eines zeitabhängigen zyklischen Attributs gegenüber dem Wert in der Vorversion nicht unbedingt ändern. Kommt es zu keiner Wertänderung, so wird jedoch die Information, daß es keine Änderung gegeben hat (obwohl prinzipiell eine Wertänderung möglich gewesen wäre), ebenfalls abgebildet. Ein Beispiel für ein zeitabhängiges zyklisches Attribut wäre etwa der Gehalt eines Mitarbeiters. Bekommt ein Mitarbeiter keine Gehaltserhöhung, obwohl prinzipiell eine Gehaltserhöhung möglich gewesen wäre, so bleibt der Gehalt konstant, die Information über die nicht erfolgte Gehaltserhöhung wird aber ebenfalls abgebildet.

Zeitabhängige Attribute im weiteren Sinn, zeitabhängige Attribute im engeren Sinn und zeitabhängige zyklische Attribute können unter dem Überbegriff zeitabhängige Attribute zusammengefaßt werden.

Die Festlegung, welche Attribute zu welcher Kategorie gehören, liegt in den Händen des "Datenbankdesigners". Das bedeutet, daß schon in der Phase des konzeptionellen Datenbankentwurfs eine Entscheidung darüber getroffen werden muß, welche Attribute welche Bedeutung haben.

Eine Datenbank mit zeitbezogenen Daten muß nun so entworfen werden, daß in einer zeitbezogenen Relation neben den Zeitstempelattributen nur Attribute

enthalten sind, die derselben Attributklasse (Klasse 1 bis Klasse 4) angehören und nicht voneinander temporal abhängig sind. Die Definition der temporalen Abhängigkeit von Attributen wird dabei vom TRM ((Navathe and Ahmed 89) und (Navathe and Ahmed 93)) übernommen.

Im Vergleich mit den im Abschnitt 3 dieser Arbeit beschriebenen Modellen hat dieser Ansatz den Vorteil, daß auf den Entwurf von Datenbanken mit zeitbezogenen Daten und damit auf die Integritätsbedingungen explizit eingegangen wird. Das Coalescing-Problem wird mit der Einführung der unterschiedlichen Attributarten wesentlich entschärft. Durch die Abbildung im relationalen Modell können so zeitbezogene Daten in verfügbaren Datenbankmanagementsystemen abgebildet werden.

6. Zusammenfassung und Schlußfolgerungen

Im vorliegenden Beitrag wurde versucht, die Entwicklungen der letzten Jahre auf dem Gebiet der temporalen Datenbanken zu skizzieren. Es hat sich dabei gezeigt, daß es trotz einer großen Anzahl von temporalen Datenmodellen und Entwürfen für temporale Datenbanksprachen noch immer kein verfügbares temporales Datenbankmanagementsystem gibt. Ein solches ist auch mittelfristig nicht zu erwarten. Ansätze, die versuchen, die Zeitdimension in dem in vielen DBMS verfügbaren relationalen Datenmodell sauber zu integrieren, erhalten daher eine immer größere Bedeutung. Die Eckpunkte eines solchen Ansatzes wurden in dieser Arbeit vorgestellt, wobei insbesondere auf den sonst in der Literatur weitgehend vernachlässigten Aspekt des Datenbankentwurfs Rücksicht genommen wurde. Die weitere Umsetzung, Erweiterung und Implementierung dieses Ansatzes wird Inhalt zukünftiger Forschungsarbeiten und Projekte sein. Dabei wird in einiger Zeit auch auf den zeitbezogenen Teil des SQL3-Standards zurückgegriffen werden können.

Literatur

[Allen 83]

J.F.Allen. *Maintaining knowledge about temporal intervals*. Communications of the ACM, 1983, S. 832-843

[Bergamaschi and Sartori 98]

S. Bergamaschi and C.Sartori. *Chrono: a conceptual design framework for temporal entities*. Proceedings of the 17th International Conference on Conceptual Modeling (ER'98), 1998

[Böhlen and Jensen 96]

M. Böhlen and C.S. Jensen. *Seamless integration of time into SQL*. Technical report, R962049, Aalborg University, Department of Computer Science, 1996.

[ISO/IEC 97]

ANSI TC X3H2, ISO/IEC JTC 1/SC 21/WG 3. *ISO Working Draft Temporal (SQL/Temporal)*, October 1997

- [Kaiser 96]
A. Kaiser. *Zeitbezogene Datenbanksysteme: eine Bestandsaufnahme und ausgewählte Problemstellungen*. Proceedings 5. Internationales Symposium für Informationswissenschaft ISI'96, 1996, S. 143-155
- [Kaiser 98a]
A. Kaiser. *Referential integrity constraints in temporal relational databases*. Working paper, submitted for publication, 1998.
- [Kaiser 98b]
A. Kaiser. *Überlegungen zur Entitätsintegrität in temporalen relationalen Datenbanksystemen*. Proceedings des 4.ZOBIS-Workshops Hamburg. 1998.
- [Lorentzos 93]
N.A.Lorentzos. *The Interval Extended Relational Model and its Application to Valid Time Databases*. J.Clifford and R.Snodgrass(eds.), Temporal Databases: Theory, Design and Implementation, 1993, S. 67-91
- [Lorentzos and Mitsopoulos 97]
N.A. Lorentzos and Y.G. Mitsopoulos. *SQL Extension for Interval Data*. IEEE Transactions on Knowledge and Data Engineering, 1997, S.480-499
- [Myrach 97]
T. Myrach. *TSQL2: Der Konsens über eine temporale Datenbanksprache*. Informatik-Spektrum, 1997, S.143-150
- [Navathe and Ahmed 89]
S. Navathe and R. Ahmed. *A Temporal Relational Model and a Query Language*. Information Sciences, 1989, S.147-175
- [Navathe and Ahmed 93]
S. Navathe and R. Ahmed. *Temporal extensions to the relational model and SQL*. In Tansel,A.: Temporal databases, Benjamin Cummings, 1993, S.92-109
- [Snodgrass et al. 95]
R. Snodgrass et al. *The TSQL2 Temporal Language*. Kluwer Academic Publishers, 1995.
- [Steiner 98]
A. Steiner. *A Generalisation Approach to Temporal Data Models and their Implementations*. Dissertation ETH Zürich, 1998
- [Tansel 93]
A. Tansel. *SQL^T: A Temporal Extension to SQL*. In Snodgrass,R.(Ed.), Proceedings of the International Workshop of an Infrastructure for Temporal Databases, 1993, S. II1-II14
- [Tansel 97a]
A. Tansel. *The Expressive Power of Temporal Relational Query Languages*. IEEE Transactions on Knowledge and Data Engineering, 1997, S.120-134
- [Tansel 97b]
A. Tansel. *Temporal relational data model*. IEEE Transactions on Knowledge and Data Engineering, 1997, S. 464-479
- [TimeDB 98]
A TimeConsult Product. *TimeDB 2.0 Demo Version*, <http://www.TimeConsult.com>, 1998

Cultural Aspects of Group Support Systems

Fahri Yetim

Marmara University Istanbul
German Dept. of Information Systems
Anadoluhisari Kampüsü, TR-81616 Anadoluhisari - Istanbul
E-Mail: yetim@hisar.aenfor.marun.edu.tr

Summary

This paper is concerned with cultural factors in Group Support Systems (GSS) and the impact of technology and culture on the process and outcome of group decision making. Both current research findings in various cultural settings and propositions for further theoretical and empirical investigation are presented.

Contents

1. Introduction
2. Conceptualizing cultural differences
 - 2.1. Dimensions of culture: a brief overview
 - 2.2. Hofstede's dimensions of culture
 - 2.3. Limitations of cultural dimensions
3. Culture and GSS
 - 3.1 GSS vs. face-to-face communication
 - 3.2 Adding culture to GSS: research findings
 - 3.2.1. Consensus change
 - 3.2.2. Group polarization and persuasiveness
 - 3.2.3. System usage
 - 3.2.4. Participation
 - 3.2.5. Communication between diverse groups
4. Conclusion

1. Introduction

As a result of modern technology and global interdependence, communication and cooperation between individuals from the same as well as from different cultures is occurring increasingly. The need to provide better tools for supporting groups shifts our attention from how our communication and computing technologies work to how people use these tools in their work (Winograd 1997) and, furthermore, how people use tools together are cultural matters (Yetim 1998).

In recent years, a few studies concerning the interaction of culture and information technology have been conducted in different research fields such as Computer-Supported Cooperative Work (Ishii 1993, Connolly 1994, Yetim 1995a/b) and Human-Computer Interaction (Yeo 1996) as well as Hypertext Technology (Yetim 1996). Researchers have suggested that communication media can change group behavior (Nunamaker et al. 1991) and culture is an important determinant of group decision making (Hofstede 1981, Sauter 1992, Watson et al. 1994). These suggestions led researchers to examine explicitly culture as an influencing factor on both the outcome and the process of decision making. Technology used includes Group Support Systems (GSS), which combine computer, communication, and decision technologies to support problem formulation and solution in group settings (Nunamaker et al. 1997). In this paper, I will discuss current research efforts as well as some propositions concerning cultural aspects of GSS by considering both the understanding of culture and the findings based on it.

2. Conceptualizing cultural differences

2.1. Dimensions of culture: a brief overview

In the literature, there are many definitions of the term culture, which emphasize either the behavioral aspects of people's lives at a point in time, or the psychological ones. Some studies dimensionalize culture by dealing with various values. Kluckhohn and Strodtbeck (1961) have identified five dimensions of cultural differences such as time orientation, person-nature orientation, activity orientation, human nature perception, and relational orientation. Hofstede (1981, 1991) has identified dimensions of national culture such as individualism vs. collectivism, high vs. low power distance, high vs. low uncertainty avoidance, and masculinity vs. femininity. These have been widely applied by other researchers. Somewhat similar are Cohen's (1987) dimensions such as indirectness vs. directness, exaggeration vs. understatement, high vs. low impersonalism, and formality vs. informality. More recently Trompenaars and Hampden-Turner (1998) argued that every culture distinguishes itself from others by specific solutions it chooses to certain problems which (1) arise from our relationships with other people, (2) come from the passage of time; and (3) relate to the environment. They identified seven fundamental dimensions of

culture. Five of them concern the relationships with people such as universalism vs. particularism, individualism vs. communitarianism, neutral vs. emotional, specific vs. diffuse, and achievement vs. ascription. The other two dimensions of culture are the attitudes to time and attitudes to the environment. It is beyond the scope of this article to describe all these dimensions in details. However, a brief explanation of the Hofstede's dimensions is given, which have been used successfully by most GSS researchers.

2.2. Hofstede's cultural dimensions

Power distance is defined as „the extent to which the less powerful members of institutions and organizations within a country expect and accept that power is distributed unequally (Hofstede 1991). In high-power-distance cultures, there are considerable differences between low- and high-status individuals. Subordinates in such cultures tend to defer to superiors and do not question their authority. In low-power-distance cultures, inequalities among people are minimized.

Uncertainty avoidance is defined as the degree to which uncertainty and unpredictability are tolerated in a society. In societies with a high degree of uncertainty avoidance, people feel uncomfortable in unstructured or risky situations. Rules and procedures are more necessary for employees in cultures with high uncertainty avoidance.

Individualism vs. collectivism refers to the relationship between the individual and his or her larger environment. In individualistic cultures, ties between individuals are loose, and the interests of individuals and individual rewards and recognition prevail over those of the group. On the other hand, in collectivistic cultures, group interests tend to prevail over individual recognition and rewards. The focus in collectivism is on group harmony rather than on confrontation. When making group decisions, relationships are considered more important than the task at hand.

Masculinity vs. Femininity describes the relative tradeoff between an assertive environment vs. supportive environment. Cultures with high masculinity scores emphasize power, assertiveness, and individual achievement. Performance-contingent rewards is consistent with such a culture. However, cultures with high femininity scores place a high value on the quality of life and caring for others. Interpersonal relationships and quality of work-life issues are consistent with femininity.

2.3. Limitations of cultural dimensions

The recent debate over how to conceptualize cultural differences challenge the study of group differences as single bipolar constructs (Kim et al. 1997, Wink 1997). The danger of stereotyping are beginning to be expressed. In the previous research, individualism and collectivism are defined in terms of one another, and cultures high in collectivism are assumed, explicitly or implicitly, to be low in individualism and vice versa. This „either-or“ assumption of the bipolar model has been challenged by a coexistence model of individualism and collectivism. Although proponents of the coexistence model tend not to

question the fact that Asian cultures are more collectivistic than Western ones, they contest the view that high collectivism precludes high individualism. It is argued, for example, that Asians accommodate the importance of both group and individual goals, or interdependence and independence of the self, by embracing and tolerating contradictions (Sinha/Tripathi 1994), by maintaining a distinction between the public and the private self, and by drawing a distinction between in-group and out-group behavior. Hofstede's model has also been criticized because it is based on a study of 116,000 respondents taken exclusively from one multinational company (i.e., IBM). It is questioned whether the personnel of a single multinational corporation are representative of a particular culture. In addition, since Hofstede's data were collected in the 1970s, there may have been significant changes in the cultures studied. Finally, it should also be remarked that studies on dimensions of culture concern only general cultural features and do not take into consideration the necessary interaction that characterizes all interpersonal or group communication: it is not enough to describe the participants separately - the model must show them together, which implies negotiating (or imposing) a common ground. These remarks do not question the usefulness of these dimensions for studying cultural factors of technology. However, they can help to view the research findings or propositions carefully.

3. Culture and GSS

In the literature of Group Support Systems (GSS), which have been designed to support group work, much of the relevant research attempts to compare group process and outcomes with and without technology. Recently, scholars have begun to address not only the question of what are the influences of media on decision making, but also the issue of how these influences are a function of culture.

3.1. GSS vs. face-to-face communication

GSS research has reported differences in group decision outcomes when comparing GSS meetings with face-to-face (FTF) meetings (Nunamaker et al. 1991). The difference in the outcomes result from the significant differences in the process of decision making provided by the GSS environment. FTF group meetings are seen as inefficient and ineffective because, as some studies have shown, individuals in a group setting may be apprehensive about how their ideas will be received, or some members of the groups dominate the discussion, causing premature consensus as well as a reduction in the collective values of the discussion. GSS therefore is seen as a technology that facilitates group interaction by empowering individuals to be more productive group members. Further, many GSS researchers argue that provision of an electronic communication channel enhances information exchange within a group and leads to a more balanced involvement of group members, which will in turn lead to better decision outcomes (for further benefits of GSS usage see Nunamaker et al. 1991/1997 and El-Shinnawy/Vinze 1998).

The features of GSS most closely associated with these benefits are *anonymity* and *parallel communication*. Anonymity reduces evaluation apprehension, conformance pressure and social cues. It is argued that a reduction in any of these „may encourage a more open, honest and free-wheeling discussion of key issues“ (Nunamaker et al. 1991, 55). Anonymity can reduce also the domination of some group members and allows participation based on content of communication rather than on the source that generated the communication. The parallel communication feature of the GSS gives group members the opportunity to contribute and work simultaneously. Good ideas or important thoughts are captured in a GSS setting whereas in a FTF meeting they are frequently lost because only one person can talk at a time. However, both anonymity and parallel communication can also have a negative impact on the group process. For example, anonymity can increase free riding, while parallel communication can cause information overload.

3.2. Adding culture to GSS: research findings

Watson et al. (1994) pointed out, arguments concerning the usefulness of communication media based on implicit assumptions that may be culturally specific: For example, an assumption is made that it is important for each group member to have an equal opportunity, regardless of status differentials, to express an opinion in a group discussion. It has also assumed that group members prefer open and direct communication to resolve conflict or disagreement. In addition, the GSS concept implicitly suggests that group decisions should further organizational objectives and little consideration is given to group harmony. If these assumptions are not valid for all cultures, they may not be acceptable, and existing GSS designs may not be appropriate for all cultures. Furthermore, if the arguments are true that different cultures require different kinds of information, process information differently, and have different degrees of satisfaction with information systems (Watson et al. 1994), then current theories which support GSS research and drive its design, as well as the majority of empirical GSS research using local subjects in Western culture, may not hold in all cultures.

A few studies have explicitly investigated cultural factors in GSS research, i.e. the impact of technology and culture and their interaction on the process and outcome of group decision making. The described studies examine different phenomena in GSS environment such as consensus change, group polarization and persuasiveness of arguments, the use of systems, and participation. To measure the impact of culture on these phenomena, the majority of these studies employ Hofstede's dimensions of culture.

3.2.1. Consensus change

- *Impact of individualism and power distance on consensus change:*

Watson et al. (1994) investigated the impact of culture on the change in consensus resulting from GSS use. They used experimental data collected in Singapore and in the US, which culturally differ with respect to individualism and power distance. Consensus change was determined by comparing

consensus before and after the meeting. Singaporean groups were found to have higher pre-meeting consensus and less change in consensus than US groups. This is consistent from a cultural perspective because Singaporean groups are more likely to be collectivists than US groups and collectivists more highly value and strive for consensus independent of GSS use. Therefore the ability of GSS to increase consensus will be diminished when employed by collectivists. Similarly, Mejias et al. (1996) found that Mexican FTF groups obtained higher consensus than GSS-supported groups, though this was not the case in the US. This finding is reasonable given the more collectivist culture of Mexico.

3.2.2. Group polarization and persuasiveness

El-Shinnawy and Vinze (1997) examined the phenomenon of group polarization - the tendency of individuals in a group setting to engage in more extreme decisions than their original private individual decisions. They define polarization as a choice shift resulting from group discussion. Using the theory of persuasive arguments which asserts that the degree of persuasiveness of arguments during group discussion can have a major effect on group polarization, they argued that polarization as a phenomenon and persuasive arguments as an explanation for it are largely US based. Persuasive arguments theory assumes that individuals are driven to present their ideas and influence others by their need to promote self-interest and to strive for self-actualization. This theoretical perspective is based on notions of individualism. Studying cultural settings that promote collectivism and accept higher degrees of power distance, would lead to low levels of polarization in the group's decision and persuasive arguments.

They compared decision making processes and outcomes of both FTF and GSS groups in the US and Singapore. Their results confirm that GSS technology does indeed have an impact on group decision outcomes such as polarization. However, their findings show a more interesting effect on the process of group decision making.

- *Impact of individualism and power distance on persuasive arguments:*

The effect of GSS on persuasiveness was most pronounced in the US groups. The US FTF groups have the highest level of persuasiveness in an argument pool resulting from group discussion, while the US GSS groups have the lowest level of persuasiveness. El-Shinnawy and Vinze (1997) explain the high level of persuasive content in the argument pool by the cultural traits of high individualism and low power distance attributed to groups operating in the US. In such cultures, group members are expected to be most active in contributing to the group discussion. The presence of higher-status individuals did not deter group members from contributing ideas that they consider important. On the other hand, Singaporean FTF groups experienced much lower levels of persuasive content in the argument pool due to the collective and high power distance of the culture, which suggests that contributions to the argument pool were based on support for the powerful person in the group rather than on the validity or novelty of the argument itself. Validity and novelty have been

considered as the two important determinants of the persuasiveness of an argument pool.

Contrary to the expectations that GSS would increase the total number of arguments and thus the level of persuasiveness of the argument pool, GSS dramatically reduced persuasive arguments in the US setting and had very little impact in the Singaporean setting. The dramatic effect of GSS in reducing the persuasive content of arguments in the US is accentuated by the initial high level of persuasive content of the argument pool for the US FTF groups. Indeed, they recognized a higher number of arguments in the GSS session than in the FTF session. But, their persuasiveness was lower. According to El-Shinnawy and Vinze, these situations occur because parallel communication facility led to redundant comments, that do not contribute to the persuasiveness of the argument pool. They are neither new ideas nor new ways to address the task at hand. In addition, anonymity did not allow individuals to contribute more novel and creative ideas through reduced inhibition in the US groups. Free riding, as a possible result of anonymous communication, is a possible explanation for it: The high level of individualism that characterize the US culture may have caused group members to abstain from making their most notable contributions in a GSS setting because in this medium the source of the ideas is not identifiable and, therefore, individuals making these ideas would not get credit for the idea.

Also against their expectation, GSS had very little impact on persuasiveness in Singaporean groups. Where conformity is the norm, such as is the case in Singapore, their expectation was that individuals would not contribute novel arguments in a FTF setting. In technology-supported groups where individuals are anonymous and parallel communication is possible, they expected, to see a significant difference in the novelty of the pool of arguments. El-Shinnawy and Vinze believe that the collectivist and high-power-distance attributes of the Singaporean cultures are so strong that the presence of technology did little to influence individuals' inherent need to conform and lack of motivation to be distinct and different from others.

3.2.3. System usage

In order to understand the use of information systems, the Technology Acceptance Model (TAM) has been widely used by researchers. TAM explains the behavior of computer usage and posits that the acceptance and use of information technology are affected by both perceived usefulness, i.e. the extent of people's belief that the technology will help them to perform their job better, and perceived ease-of-use, i.e. the degree to which a person believes that using a particular system would be free of effort (Davis et al. 1989). To understand whether TAM applies in other cultures, Straub et al. (1997) tested TAM model across three different countries such as Japan Switzerland, and the US, by employing E-mail technology as an example.

- *Impact of uncertainty avoidance, power distance, individualism, and masculinity on technology acceptance:*

Straub et al. (1997) predicted that cultural differences in *uncertainty avoidance* could affect technology acceptance by influencing choices of computer-based media vs. traditional media. Based on the assumptions of both information richness and social presence theories that the individuals choose media on the basis of how well those media reduce uncertainty, individuals in high uncertainty avoidance are expected to use electronic media less often since these media are not well suited to uncertainty reduction as face-to-face and other rich channels. *Power distance* should have a marked bearing on the communication patterns of knowledge workers and their managers. In societies in which managers and workers are separated by a large power distance, the leveling effect of computer-based media is not seen or felt as a desirable feature. In high power distance cultures, individuals may show deference to authority by refraining from using media that do not allow them face-to-face contact. Concerning the effect of *individualism* vs. *collectivism*, they predicted that knowledge workers in collectivist cultures cannot pick up cues about the social situation as readily from computer-based media and, therefore, would be inclined, overall, toward media such as face-to-face across all communications tasks. Because *Masculinity* (assertiveness) is highly related to interpersonal presence, media not conveying the social presence of the communicator, such as E-mail, would not be favored in cultures with high masculinity. Their results indicate that TAM holds for both the US and Switzerland, but not for Japan, suggesting that the model may not predict technology use across cultures. In case of Japan, cultural tendencies toward more uncertainty avoidance, greater power distances between managers and workers, collectivist sentiments, and assertiveness may limit E-mail use and disassociate usefulness from use. Because the study did not directly measure cultural dimensions, "it is not possible to say with certainty that a link between cultural factors and technology has been empirically established." (Straub et al. 1997, 9)

3.2.4. Participation

Group participation is also affected by social norms. Therefore, the ability of GSS to reduce participation barriers may depend on culture in a given context. By explaining GSS participation in terms of the technology acceptance model (TAM), Robichaux/Cooper (1988) considered some determinants of ease of

Participation barriers	Power distance	Individualism-collectivism	Masculinity-femininity
<i>Parallel entry</i>			
Time domination	Increases participation of low-power members in high-power distance cultures;	Increases participation of non-in-group members in collectivist cultures;	Increases participation;
Dispensive free riding	Decreases participation of low-power members in high-power distance cultures;		
Media incompatibility	Decreases participation of high-power members in high-power distance cultures;		
Conflict avoidance		Decreases participation in collectivist cultures;	
Indifference free riding			Decreases participation in masculine cultures;
<i>Anonymous entry</i>			
Time domination	Increases participation of low-power members in high-power distance cultures;	Increases participation of non-in-group members in collectivist cultures;	
Evaluation apprehension	Increases participation of low-power members in high-power distance cultures;		
Conformance pressure	Increases participation of low-power members in high-power distance cultures;		
Dispensive free riding	Decreases participation of low-power members in high-power distance cultures;		
Conflict avoidance		Increases participation in collectivist cultures;	Increases participation in feminine cultures;
Indifference free riding		Decreases participation in individualist cultures;	Decreases participation in masculine cultures;

Fig 1: Propositions for GSS-culture interaction (cf. Robichaux/Cooper 1998, 292)

participation and usefulness of participation such as

- *Time domination* occurs when an information channel is dominated by members with some skills
- *Media incompatibility* occurs when user has not the ability to use a given media

- *Indifference free riding* occurs when a member believes there is no personal value participating and thus does not participate
- *Dispersive free riding* occurs when a member believes that other members have the same or better ideas and therefore does not participate
- *Evaluation apprehension* occurs when a member who wants to avoid negative evaluation does not participate
- *Conformance pressure* occurs when a member who wants to avoid punishment from powerful members does not participate
- *Conflict avoidance* occurs when a member does not participate in order to avoid conflict among members, and thereby enhance group maintenance

These can affect the value of participation. By employing Hofstede's dimensions of culture as example settings, Robichaux and Cooper (1998) hypothesized some participation barriers as affected by GSS features such as parallel and anonymous entry (fig. 1). However, as Robichaux and Cooper (1998) remarked, these propositions should be viewed as highly tentative due to the Western bias of the technology acceptance model, which the propositions are based upon. This model itself focuses on an individual's decisions to participate based on cognitive evaluations of his or her beliefs and perceptions. It could be contrarily argued that emotion rather than rational reasoning can predominate in some cultures. But, the authors assume that rational decision making and GSS behavior, such as participation, are applicable across-cultures.

3.2.5. Communication between diverse groups

The studies presented thus far were based on an assumption that individuals of a single culture were employing GSS. They made a significant contribution to the cultural aspects of GSS. However, they did not consider the impact of GSS when participants represent multiple cultures. This issue is getting more important in the age of globalization. Cultural differences in terms of language can affect the interaction between personal abilities and the supporting technologies, which result in a differentiation of behavioral skills and lead to differences in the ease of use and thus affect the participation in communication. Beside language barriers, culturally diverse groups may also be hindered by behavioral constraints posed by different cultural norms as well as discrimination. These groups might be assisted by computer-based systems that facilitate group interaction, and decision making by allowing individual idea generation, anonymity, parallel communication, providing multilingual interfaces that automatically translate comments to an individual's native language (Aiken et al. 1994, Yetim 1995a,b).

In a study, Daily et al. (1996) compared the effects of using group decision support systems on culturally diverse and homogeneous groups. Group performance was measured by the number of non-redundant, realistic ideas and the quality of solution produced in response to posed problems. Groups using the system, regardless of cultural make up, developed a significantly higher number of non-redundant, realistic ideas than groups that did not. Furthermore, among groups using the system, culturally diverse groups

produced a significantly higher number of non-redundant, realistic ideas than homogeneous groups. But, they could not find any significant indication whether culturally diverse groups that use a group support system will generate a higher quality solution to a posed problem than culturally homogeneous groups.

Another empirical study concerning the impact of technology on communication among culturally diverse groups have been conducted by Ma (1996). After studying E-mail communications between East Asians and North American students, Ma found support for the hypotheses that both groups tend to be more direct and show greater self-disclosure in computer-mediated conversations than in FTF conversations and both groups are by communicating through media less likely to adapt to each other's cultural rules than FTF communication. The latter is based on the reasoning that the concept of „stranger“, which characterizes FTF intercultural communication situations and also implies that there is a host/guest distinction does not apply. Either all are „strangers“ or none is a „stranger“. They are not bound by any particular set of cultural rules.

However, these assumption can be questioned with the argument that knowing the place or culture of communication partner could still have an impact on communication behaviour in computer-mediated communication environment. Thus, I suppose that

- *Intercultural computer-mediated communication between two culturally different groups in one society will differ from communication between the same groups in another society (hypothesis 1); and*
- *Chosen language can have an impact on the relationship between high and low-status members in a group, because different languages offer different way to express relationships (hypothesis 2).*

The test of these hypotheses that have not been considered in the intercultural computer-mediated communication thus far, could inform us about the interaction between technology, language, culture, as well as the place of communication.

4. Conclusion

These research results confirm that technology and culture and their interaction indeed have an impact on group processes and their outcome. They also confirm that the extent of the impact is quite different depending on the culture that dictates the norms under which a group operates.

These results have implications for both GSS developers and researchers. For developers and designers of GSS technology, they provide impetus for incorporating cultural factors into existing GSS applications. For researcher, they signal the need to identify further cultural factors that might potentially hinder implementation of these systems in culturally different settings. A

challenge to managers implementing GSS, for example, in high-power distance and collectivist cultures is to reduce the barriers. As Robichaux and Cooper (1998) discussed, if ideas are desired from low-power members specifically, meetings of only low-power members could be held to reduce dispersive free riding. Since members may withhold comments or ideas that oppose those of others to avoid conflict, managers could have members initially generate but not exchange ideas. This may reduce the knowledge that members have of one another's ideas and thus reduce instances of their withholding ideas that they feel are in conflict.

However, these studies have some limitations, thus, implications drawn from this research should be viewed with caution. First, research findings are inconsistent concerning the effects of a GSS. This is not surprising because many of the cited studies differ on a variety of important factors. Some studies suggest that a GSS may increase group consensus, some studies report no change. Second, most studies are based only on general features of culture and on the „either-or“ assumption of the bipolar model which has been challenged by a coexistence model of individualism and collectivism. Third, the studies have some experimental limitations. Some of them are based on empirical data which „were not collected under experimentally controlled conditions“ (Ma 1996, 184). Others used only a small sample size for the study and did not test hypotheses on different group support systems (Daily et al. 1996, 287). Different features provided by different GSS may affect group interactions to greater or lesser degree than the systems they used. Some of them did not measure cultural dimensions (Straub et al. 1997, Ma 1996). Thus, it is not possible to say with certainty that a link between cultural factors and technology has been empirically established. The propositions generated by Robichaux and Cooper (1998), but not tested empirically, should be viewed as highly tentative, and regarded as opportunities for further theoretical and empirical investigation.

Further research is needed, especially concerning the cooperation and communication of diverse groups. Designing modern communication systems that will help people overcome the cultural barriers to communication and cooperation is not only a linguistical and sociological issue, but also a philosophical one, as discussed in (Yetim 1998), because global communication systems require to some extent the establishment of communication and cooperation norms for common orientation of participants to solve common problems and work on shared tasks. Therefore, the question is whose norm?

Acknowledgements

I would like to thank C. Porter, Dean of „Undergraduate and Interdisciplinary Studies“ at the University of California (Berkeley) for the opportunity to conduct this research. I also thank participants of the Research Colloquium at the Fisher Center of Management and Information Technology (Haas School of Business), for their helpful comments on the earlier presentation.

References

- [Aiken, M.W.; Martin, J.S.; Paolillo, J.G.P.; Shirani, A.I. 1994]: A group decision support system for multilingual groups. In: *Information & Management* 26 (1994), 155-161
- [Cohen, R. 1987]: Problems of intercultural communication in Egyptian-American diplomatic relations. In: *International Journal of Intercultural Relations* 11 (1987), 29 - 47
- [Connolly, J. H. 1994]: Artificial Intelligence and Computer Supported Cooperative Working in international Contexts. In: Connolly, J. H; Edmonds, E. A. (eds.): *CSCW and Artificial Intelligence*, Berlin et al.: Springer-Verlag, 141-159.
- [Daily, B.; Whatley, A.; Ash, S.R.; Steiner, R.L. 1996]: The effects of group decision support systems on culturally diverse and culturally homogenous group decision making. In: *Information & Management* 30 (1996), 281-289
- [Davis, F.D., Bagozzi, R.P.; Warshaw. P.R. 1989]: User acceptance of computer technology: A comparison of two theoretical models, *Management Science* 35, 1989, 982-1003
- [El-Shinnawy, M; Vinze, A.S. 1997]: Technology, Culture and Persuasiveness: a study of choice-shifts in group settings. In: *Human-Computer Studies* 47 (1997), 473-496
- [Hofstede, G. 1991]: *Cultures and Organizations: Software of the mind*. London: McGraw-Hill.
- [Hofstede, G. 1981]: *Culture's consequences: International differences in work-related values*. London: Sage.
- [Ishii, H. 1993]: Cross-Cultural Communication and CSCW. In: Harasim, L.M. (ed.): *Global Networks. Computers and International Communication*. Cambridge et al.: The Mit Press, 143-151.
- [Kim, U; Triandis, H.C.; Kagitcibasi, C.; Choi, S.-C.; Yoon, G. 1994](eds.): *Individualism and collectivism: Theory, method, and applications*. Thousand Oaks, CA: Sage.
- [Kluckhohn, F.R.; Strodtbeck, F.L. 1961]: *Variations in value orientations*. Evanston: Row Patterson and Co.
- [Ma, R. 1996]: Computer-Mediated Conversations as a New Dimension of Intercultural Communication between East Asian and North American College Students. In: Herring, S. C. (ed.): *Computer-Mediated Communication*. Amsterdam/Philadelphia: John Benjamins Publishing company. 173-185.
- [Mejias, R.J.; Shepherd, M.M.; Vogel, D.R.; Lazaneo, L. 1997]: Consensus and perceived satisfaction levels: A cross-cultural comparison of GSS and non-GSS outcomes within and between the United States and Mexico. In: *Journal of Management Information Systems* 13 (1996-1997), 137-161
- [Nunamaker, J.F.; Briggs, R.O.; Mittleman, D.D.; Vogel, D.R.; Balthazar, P.A. 1997]: Lessons from a Dozen Years of Group Support Systems Research: A Discussion of Lab and Field Findings. In: *Journal of Management Information Systems* 13 (1996-1997), 163-207
- [Nunamaker, J.F.; Dennis, A.R.; Valacich, J.S.; Vogel, D.R.; George, J.F. 1991]: Electronic meetings to support group work. In: *Communications of the ACM* 7(1991), 40-61
- [Robichaux, B.P.; Cooper, R.B. 1998]: GSS participation: A cultural examination. In: *Information & Management* 33 (1998), 287-300
- [Sauter, V. L. (1992): Cross-Cultural Aspects of Model Management Needs in a Transnational Decision Support System. In: Palvia, S.; Palvia, P.; Zigli, R. (eds.):

- The Global Issues of Information Technology Management. Idea Group Publ., 332-355.
- [Sinha, D.; Tripathi, R.C. 1994]: Individualism in a collectivist culture: A case of coexistence of opposites. In: U. Kim, H.C. Triandis, C. Kagitcibasi, S.-C. Choi; G. Yoon (eds.): Individualism and collectivism: Theory, method, and applications. Thousand Oaks, CA: Sage. 123-136.
- [Straub, D.; Keil, M.; Brenner, W. 1997]: Testing the technology acceptance model across cultures: A three country study. In: Information & Management 33 (1997), 1-11
- [Trompenaars, F.; Hampden-Turner, C. 1998]: Riding The Waves of Culture. New York et al.: McGraw-Hill. Second edition
- [Watson, R.T.; Ho, T.H.; Raman, K.S. 1994]: The Fourth Dimension of Group Support Systems. In: Communications of the ACM 10 (1994), 44- 55.
- [Wink, P. 1997]: Beyond Ethnic Differences: Contextualizing the Influence of Ethnicity on Individualism and Collectivism. In: Journal of Social Issues 2(1997), 329-350
- [Winograd, T. 1997]: The Design of Interaction. In: Denning, P.J. et al. (eds.): Beyond Calculation: The next Fifty Years of Computing. New York: Copernicus, 149-161
- [Yetim, F. 1998]: Interkulturalität und Informatische Gestaltung. Eine interdisziplinäre Annäherung. In: Informatik-Spektrum 4 (1998), 206-215
- [Yetim, F. 1996]: Multilinguale Aspekte von Text und Hypertext. In: Krause, J.; Herfurth, M.; Marx, J. (Hrsg.): Herausforderungen an die Informationswirtschaft. Informationsverdichtung, Informationsbewertung und Datenvisualisierung. Konstanz: Universitätsverlag, 109-120
- [Yetim, F. 1995a]: Interkulturelle Kommunikationsprobleme und die Konsequenzen für CSCW. Eine neue Betrachtung des Sprach-Handlungs-Modells von Winograd und Flores. Künstliche Intelligenz 2 (1995), 45-47
- [Yetim, F. 1995b]: Internationalisierung der Telekommunikation: Der Kick für interkulturelle Kommunikation und Kooperation? In: Schieber, P. (Hrsg.): Informationsmanagement in der Informationsgesellschaft, Konstanz: Universitätsverlag, 287-301
- [Yeo, A. 1996]: Cultural User Interfaces. A Silver Lining in Cultural Diversity. In: SIGCHI Bulletin 3(1996), 4-7

Das Konzept einer strikt benutzerorientierten Evaluierung maschineller Übersetzungssysteme (am Beispiel des Projekts MIROSLAV)

Jutta Marx, Bernhard Staudinger

Dr. Jutta Marx
Dr. Bernhard Staudinger
FG Informationswissenschaft
Universität Regensburg
93040 Regensburg
e-mail: jutta.marx@sprachlit.uni-regensburg.de

Inhalt

- 1 Vorbemerkung
- 2 Das Projekt MIROSLAV
- 3 Der MIROSLAV-Anwendungspartner
- 4 Das MIROSLAV-Evaluierungskonzept
 - 4.1 Benutzerorientierte Evaluierung
 - 4.2 Einzelevaluierung
 - 4.3 Relevanz- und Informativübersetzung
 - 4.4 Übersetzungsqualität und -produktivität
 - 4.5 Globaler Evaluierungsprozeß
 - 4.6 Black-Box-Ansatz + Lexikonfenster
 - 4.7 Häufigkeit der Evaluierung
 - 4.8 Evaluierungsziele: Qualität
 - 4.9 Evaluierungsziele: Produktivität
- 5 Zusammenfassende Übersicht über die MIROSLAV-Testverfahren
- 6 Ausblick
- 7 Literatur

1 Vorbemerkung

Maschinelle Übersetzungssysteme rücken in letzter Zeit wieder vermehrt in das Interesse von Wissenschaft und Industrie. Nachdem sie lange Zeit als unbrauchbar galten, finden sie mehr und mehr Absatz, nicht zuletzt aufgrund der

Tatsache, daß Entwickler und Vermarkter abrücken von dem Anspruch, perfekte Reinübersetzungen liefern zu können, die mit Humanübersetzungen vergleichbar sind. Stattdessen enthalten maschinelle Übersetzungssysteme mittlerweile ganz selbstverständlich Posteditionshilfen, die eine komfortable Überarbeitung der von der Maschine gelieferten Rohübersetzung erlauben. Für den Benutzer bleibt (wie immer) das Problem, aus der Gesamtpalette das für seine Bedürfnisse passende Produkt auszuwählen. Die Beurteilung der Qualität eines Systems gestaltet sich dabei jedoch als ausgesprochen schwierig, da eine Vielfalt von beeinflussenden Parametern wie Textart, Verwendungszweck, Zeitdruck u.ä. eine allgemeingültige Einschätzung eines Systems unmöglich macht. Evaluationen des Vermarkters sind dabei ebenso eingeschränkt nutzbar wie die der Entwickler, da erstere natürlich bemüht sind, die Vorteile ihres Systems herauszustreichen, letztere hauptsächlich an der Verbesser- und Erweiterbarkeit der Software interessiert sind, die dem Benutzer in seiner konkreten Arbeitssituation wenig weiterhelfen. Evaluationen unabhängiger Dritter wie z.B. Computerzeitschriften mangelt es dagegen an Konkretheit. Da sie nicht auf den speziellen Anwendungskontext eines potentiellen Benutzers eingehen können, sind ihre Bewertungen höchstens als erste Orientierungshilfe zu verstehen.

Das vorliegende Papier versucht diese Lücke zu schließen, indem es ein Konzept einer benutzerorientierten Evaluation für maschinelle Übersetzungssysteme vorstellt, das das zu beurteilende Übersetzungssystem aus der Sicht eines potentiellen Benutzers untersucht. Es wird somit versucht, konkrete Aussagen über die Angemessenheit und den Nutzen eben dieses Systems für den Benutzer in seiner speziellen Arbeitssituation zu machen. Generelle Beurteilungen zum Leistungsspektrum eines Systems bzw. zu seiner Erweiter- oder Verbesserbarkeit, d.h. diagnostische Aussagen interessieren in diesem Zusammenhang weniger.

Die gestellten Aufgaben sind Teil des Projektes MIROSLAV (Machine Translation Initiative for Russian and other Slavic Languages).

2 Das Projekt MIROSLAV

MIROSLAV ist ein von BMBF gefördertes Verbundprojekt zwischen der Gesellschaft für multilinguale Systeme (GMS) in Berlin, Sietec/München, dem Institut für Slavistik der Humboldt-Universität zu Berlin und der FG Informationswissenschaft der Universität Regensburg (IWR). Hauptziel von MIROSLAV ist zunächst die Entwicklung eines maschinellen Übersetzungssystems für das Sprachpaar Russisch - Deutsch; andere slavische Sprachen sollen folgen.

Die Aufgaben der IWR bestehen dabei in der Erarbeitung von lexikalisch-semantischen Lösungen für komplexe Transfers und der Evaluation des russisch-deutschen Übersetzungssystems, das momentan als Unix-basierter Prototyp auf der Basis des Forschungssystems METAL existiert. Für Herbst 1998 ist ein PC-System angekündigt, das im Rahmen der T1-Linie von Langenscheidt vermarktet wird.

Als Anwendungspartner für die Evaluierungsarbeiten konnte die Gesellschaft sozialwissenschaftlicher Infrastruktureinrichtungen (GESIS), insbesondere die Abteilung „Informationstransfer Osteuropa“ in Berlin (kurz: GESIS Berlin) gewonnen werden.

3 Der MIROSLAV-Anwendungspartner

Die Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen e.V. unterhält seit 1992 eine Außenstelle in Berlin, die sich hauptsächlich mit der Bestandssicherung von Ergebnissen der sozialwissenschaftlichen Forschung in der ehemaligen DDR sowie dem Informationstransfer von und nach Osteuropa zu sozialwissenschaftlichen Themen beschäftigt. Darüber hinaus berät sie in Fragen sozialwissenschaftlicher Methodik sowie zu den Möglichkeiten der Informationsgewinnung in den neuen Bundesländern und in Osteuropa.

Im Rahmen des Informationstransfers zwischen den ost- und westeuropäischen Ländern führt die GESIS Berlin regelmäßige Erhebungen zu Instituten, Zeitschriften und Projekten in Osteuropa durch. Die hierdurch gewonnenen Daten werden durch Informationen aus Printmaterialien wie z.B. Zeitschriften erweitert und publiziert (z.B. im Newsletter „Sozialwissenschaften in Osteuropa“. Ein Teil dieser Informationen geht in die Bestände der vom Informationszentrum Sozialwissenschaften in Bonn (kurz: IZ Bonn) unterhaltenen Datenbank zu sozialwissenschaftlichen Forschungsprojekten (kurz: FORIS) bzw. in die von der GESIS Berlin selbst unterhaltene Datenbank zu sozialwissenschaftlichen Institutionen in Osteuropa ein (kurz: Institutionendatenbank). (Nähere Informationen zur GESIS Berlin s. Marx/Mutschke/Schommier 1995 sowie die WWW-Homepage der GESIS-Berlin unter <http://www.berlin.iz-soz.de/>.)

Zur Wahrung dieser Aufgaben fällt bei der GESIS Berlin zahlreiches Textmaterial in russischer Sprache an, das Gegenstand eines Übersetzungssystems sein könnte. Es handelt sich dabei im wesentlichen um folgende Textsorten¹:

- Im Rahmen einer bisher alle zwei Jahre durchgeführten Umfrage zu laufenden, geplanten oder bereits abgeschlossenen Forschungsprojekten in Osteuropa (d.i. Teil der sog. FORIS-Umfrage) bzgl. sozialwissenschaftlicher Problemstellungen sind von der GESIS Berlin auch zahlreiche **Fragebögen** in russischer Sprache auszuwerten.
- **Zeitschriften, Aufsätze, Monographien** und **Jahresberichte** in russischer Sprache werden regelmäßig von der GESIS Berlin in Hinblick auf Konferenzberichte, Literatur- und Forschungsprojekthinweise, Institutsbeschreibungen u.ä. ausgewertet.

¹ Die nicht maschinenlesbaren Texte wie FORIS-Fragebögen oder Zeitschriften müssen zuvor mittels Scanning und OCR erfaßt werden. Hierbei konnten mit dem Softwarepaket FineReader Professional 3.0 von ABBYY Software House bereits sehr gute Ergebnisse erzielt werden.

- Ebenfalls für *Institutsprofile*, die u.a. im Newsletter „Sozialwissenschaften in Osteuropa“ dargestellt werden oder in die Institutionendatenbank eingehen, werden **Texte im Internet** wie z.B. die Homepages osteuropäischer Institute herangezogen.

4 Das MIROSLAV-Evaluierungskonzept

Nach eingehender Sichtung der in der Literatur verfügbaren Evaluierungstypologien und -konzepte erschien das Schema nach Bourbeau 1990 am besten geeignet, die Projektziele von MIROSLAV zu repräsentieren. Das Bourbeau'sche Schema ist hierarchisch gegliedert und somit übersichtlich und erlaubt eine schrittweise Verfeinerung der angestrebten Ziele. Es umfaßt zudem die Hauptkriterien der anderen „klassischen“ Evaluierungsstrategien von Van Slype 1979, Vasconcellos 1988 und Way 1994 bzw. kann problemlos um fehlende Punkte erweitert werden.

Nachstehende Abbildung zeigt das Evaluierungskonzept in MIROSLAV, aufgebaut analog zum hierarchischen Schema nach Bourbeau, das um zwei Grundsatzentscheidungen ergänzt (d.i. benutzerorientierte Evaluierung und Einzelevaluierung) sowie um die für das Projekt wesentlichen Evaluierungsziele konkretisiert wurde.

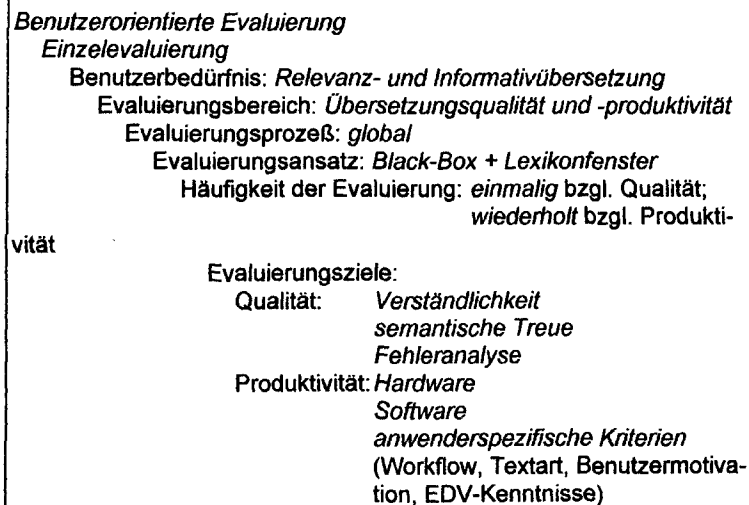


Abb. 1: Das Evaluierungskonzept in MIROSLAV (angelehnt an Bourbeau 1990:43)

Im folgenden werden nun die einzelnen Ebenen dieser Evaluierungshierarchie genauer erläutert sowie die konkreten Evaluierungsziele detailliert beschrieben.

4.1 Benutzerorientierte Evaluierung

In MIROSLAV wird eine streng benutzerorientierte Evaluierung angestrebt, die das zu beurteilende MT-System aus der Sicht eines potentiellen Benutzers untersucht². Es wird somit versucht, konkrete Aussagen über die Angemessenheit und den Nutzen eben dieses Systems für den Benutzer in seiner speziellen Arbeitssituation zu machen. Generelle Beurteilungen zum Leistungsspektrum eines Systems bzw. zu seiner Erweiter- oder Verbesserbarkeit, d.h. diagnostische Aussagen interessieren in diesem Zusammenhang weniger.

4.2 Einzelevaluierung

Im Projekt MIROSLAV wird lediglich eine Bewertung eines einzelnen MT-Systems vorgenommen. Es finden keine Vergleichstests mit anderen Systemen statt.

4.3 Relevanz- und Informativübersetzung

Statt der in der Literatur üblichen Zweiteilung des Benutzerbedürfnisses in Informativ- und Reinübersetzung wird im MIROSLAV-Evaluierungskonzept eine Dreiteilung getroffen, die zusätzlich den Verwendungszweck Relevanzübersetzung umfaßt. Hierbei ist der Benutzer lediglich darin interessiert, ob ein bestimmter Text die für seinen Verwendungszweck relevante Information enthält oder nicht, d.h. ob sich ein Text inhaltlich zur Weiterverwendung wie z.B. der Datenbankaufbereitung (s.o.) eignet³.

Neben der Relevanzübersetzung werden beim MIROSLAV-Anwendungspartner GESIS-Berlin hauptsächlich Informativübersetzungen benötigt, da die russischsprachigen Originaltexte nicht in ihrer Gänze verarbeitet werden, sondern vielmehr als Ausgangsbasis für die Pflege einer breiten Pa-

² Die Abkehr von einer generellen Evaluierung hin zu einer aufgabenorientierten Evaluierung auf dem Gebiet der Sprachverarbeitung und insbesondere der maschinellen Übersetzung setzt sich in der Community immer mehr durch (s. Rubio/Gallardo/Castro/Tejada 1998). Hier etabliert sich zur Zeit gerade der Begriff der task-orientierten Evaluierung (s. z.B. White/Taylor 1998 und Hovy 1998), die die Bewertung eines Systems in Abhängigkeit vom angestrebten Einsatzgebiet vornimmt. Benutzerorientiert ist darüber hinaus eine Evaluierung in unserem Verständnis dann, wenn sie zusätzlich den potentiellen Benutzer des zu bewertenden Systems in Betracht zieht, also z.B. auch individuelle und psychologische Kriterien (z.B. über eine Benutzermodellierung oder konkrete Benutzertests) berücksichtigt.

³ Je nach Anwendungsdomäne sind hier natürlich weitere Verwendungszwecke denkbar. So unterscheiden bspw. White/Taylor 1998 insgesamt 6 verschiedene sogenannte Text-Handling Tasks, die von Publishing (d.i. in unserer Terminologie die Reinübersetzung) über Extraction (d.i. Informativübersetzung) bis hin zu Filtering (d.i. Relevanzübersetzung) reicht.

lette von Informationsmedien (Printmedien, Datenbanken) dienen. Hierzu extrahiert man aus dem Original spezifische Informationen und bereitet sie gemäß dem Format des Zielmediums auf (s. Marx 1998a und 1998b).⁴

4.4 Übersetzungsqualität und -produktivität

Für den Nutzer eines automatischen MT-Systems ist nicht nur die Qualität der erzeugten Texte interessant, sondern auch, ob sich durch den Einsatz eines solchen Systems eine Arbeitserleichterung oder Produktivitätssteigerung erzielen läßt. Hierzu muß sich die Übersetzungskomponente gut in den Arbeitsprozeß eingliedern lassen, leicht zu bedienen sein, ggf. schnell arbeiten usw. Deshalb ist sowohl die Qualität als auch die Produktivität des MIROSLAV-Systems Gegenstand der Evaluierung in MIROSLAV.

4.5 Globaler Evaluierungsprozeß

Die Evaluierung in MIROSLAV soll anhand eines globalen Evaluierungsprozesses geschehen, d.h. daß alle verfügbaren Informationsquellen bei der Beurteilung der Übersetzungsqualität und -produktivität heranzuziehen sind. So werden nicht nur Herstellerangaben, Handbücher u.ä. zum System konsultiert, sondern auch Benutzerbefragungen und Systemtests mit potentiellen Benutzern durchgeführt.

4.6 Black-Box-Ansatz + Lexikonfenster

Kommerzielle Produkte erlauben dem Benutzer keinen Einblick in den Systemkern der Übersetzungskomponente, nicht zuletzt um zu verhindern, daß durch Systemmanipulation Funktionseinbußen oder unerwünschte Seiteneffekte eintreten. Eine Erweiterung oder Korrektur von Lexikoneinträgen ist dagegen jedoch üblich und auch sinnvoll.

Da die MIROSLAV-Evaluierung eine benutzerorientierte sein soll, verfolgt sie eben genau den Ansatz, den ein kommerzielles System zuläßt: Black-Box-Ansatz, d.h. ohne Einsicht in den Systemkern mit Lexikonfenster zur Anpassung des Lexikons.

4.7 Häufigkeit der Evaluierung

Zur Beurteilung der Übersetzungsqualität wird zunächst eine einmalige Durchführung der geplanten Tests angesetzt. Da von Seiten der GMS der russisch-deutsche Thesaurus des IZ Bonn in das MIROSLAV-System integriert wird, erscheint ein Lexikontuning und damit ein zweiter Testdurchlauf aus momentaner Sicht nicht erforderlich.

Bezüglich der Produktivität müssen dagegen zumindest zwei Durchgänge gestartet werden, da die Arbeitseffektivität vor und nach dem Einsatz eines ma-

⁴ Darüber hinaus werden in stark begrenztem Umfang auch Übersetzungen zur Weiterverwendung benötigt, z.B. bei der Publikation eines Themenbandes.

schinellen Übersetzungssystem festzustellen ist. Machen die Ergebnisse dieser Tests eine vollständige Reorganisation der Arbeitsprozesse erforderlich, so ist gegebenenfalls ein dritter Durchlauf erforderlich.

4.8 Evaluierungsziele: Qualität

Gängige Kriterien zur Feststellung der Übersetzungsqualität eines Systems sind im allgemeinen die Lesbarkeit, Verständlichkeit und die semantische Treue⁵ des Zieltextes, sowie die Art und Anzahl der im Output auftretenden Fehler. Da die Lesbarkeit eines Textes nichts über dessen Adäquatheit in Bezug auf den Quelltext aussagt (ein gut lesbarer Text kann inhaltlich falsch sein), wird sie bei der MIROSLAV-Evaluierung nicht in Betracht gezogen. Um ein möglichst umfassendes Bild der Übersetzungsqualität des MIROSLAV-Systems zu gewinnen, sollen jedoch die restlichen Kriterien in die Evaluierung einbezogen werden, wie folgende Aufstellung zeigt.

- **Verständlichkeit**

Um die Verständlichkeit der von METAL übersetzten Texte beurteilen zu können, bietet sich ein Rating-Verfahren an. Das Rating-Verfahren ist die am häufigsten angewandte Methode zur Messung der Verständlichkeit maschinell übersetzter Texte. Es basiert auf der Verwendung einer Skala, nach der Texteinheiten, meist Sätze gemäß ihrer Verständlichkeit einzustufen sind. Da sich die Verwendung umfangreicher Skalen in der Vergangenheit als äußerst problematisch herausgestellt hat (s. Marx 1998a), wendet MIROSLAV eine lediglich 3-stufige Skala an mit den Meßpunkten „verständlich“, „teilweise verständlich“ und „unverständlich“. Bemessungseinheit ist jeweils ein Satz bzw. eine Phrase im Falle von Überschriften, Aufzählungen etc.

- **Semantische Treue**

Zur Bestimmung der semantischen Treue, die ein Maß für die Genauigkeit und Treue einer Übersetzung ist, wird ein sogenannter Deskriptorentest durchgeführt. Hierbei handelt es sich um ein eigens für den MIROSLAV-Kontext entwickeltes Testverfahren, das sich aus den speziellen Gegebenheiten des Projektanwendungspartners ergibt. Möglich wird dieses Verfahren durch die Tatsache, daß die GESIS-Berlin einen Großteil der recherchierten Texte für die Projektdatenbank FORIS sowie die Institutionendatenbank aufbereitet (s.o.), d.h. u.a. formale und inhaltliche Schlagwörter vergibt, um die Suche nach Dokumenten zu ermöglichen. Angelegt ist das vorgeschlagene Verfahren auf der Basis eines Vergleichstestes, bei dem die Art und Anzahl

⁵ Statt „semantische Treue“ wird in der Literatur oft auch der Begriff „Informativität“ verwendet. Wegen der besonderen Bedeutung des Terminus „Information“ in der Informationswissenschaft, der hier eine stark pragmatische Ausrichtung hat und den Neuheitswert eines Datums für den Benutzer umschließt, soll auf den u.E. passenden Ausdruck „semantische Treue“ zurückgegriffen werden.

der für einen Originaltext vergebenen formalen und inhaltlichen Deskriptoren verglichen wird mit der Art und Anzahl der Deskriptoren, die anhand der maschinellen Übersetzung dieses Textes vergeben wurden. Je geringer die Unterschiede zwischen beiden Gruppen, desto größer ist die Genauigkeit und Verständlichkeit des maschinell übersetzten Textes. Die Hauptidee ist hierbei, daß die Deskriptoren jeweils für die wichtigsten, die den Textinhalt bestimmenden semantischen Einheiten vergeben werden und somit die Deskriptorenübereinstimmung zwischen Original und Übersetzung ein, wenn auch grober, Indikator für die Übereinstimmung in punkto semantische Treue gelten kann.

Außerdem macht der Deskriptorentest auch Aussagen über die Verständlichkeit eines maschinell übersetzten Textes, da nur ein (gut) verständlicher Text korrekt verschlagwortet werden kann.

- Fehleranalyse

Da, wie bereits beschrieben, die MIROSLAV-Evaluation keine diagnostische oder therapeutische Analyse zum Ziel hat, soll bei der Fehleranalyse der maschinellen Übersetzung eine eingeschränkte Fehlerklassifikation zur Anwendung kommen. Diese berücksichtigt lediglich diejenigen Fehlertypen, die massive Auswirkungen auf die Verständlichkeit des Textes haben. Bei der Festlegung der zu berücksichtigenden Fehlertypen haben wir uns an Rinsche 1993 sowie an Flanagan 1994 orientiert. Demnach ergibt sich folgendes Raster, das anhand einer kleinen Textauswahl der Anwendungsdomäne auf ihre Tragbarkeit hin überprüft wurde:

Berücksichtigt werden sollen:

1. falsche Segmentierung:
Die einzelnen Satzglieder sind nicht richtig erkannt worden bzw. Einzelwörter oder Satzteile werden den falschen Satzgliedern zugeordnet.
2. falsche Satzgrenzen:
Die Satzgrenzen sind falsch gesetzt, entweder durch „Zerstückeln“ eines Satzes oder Zusammenziehen zweier aufeinanderfolgender Sätze.
3. falsche Ellipsen:
Das System interpretiert eine Struktur fälschlicherweise als Ellipse, übersetzt z.B. eine Nominalphrase als Satz mit dem Kopulaverb *byť*.
4. Wort nicht übersetzt:
Ein Wort ist nicht im Lexikon vorhanden oder wird aus anderen Gründen (z.B. wegen phrasaler Analyse) nicht übersetzt (mit Ausnahme der nicht zu berücksichtigenden Wortklassen, s.u.).
5. falsche Wortwahl:
Ein Wort wird falsch, d.h. nicht kontextgemäß übersetzt.

Diese Hauptfehlerklassen sind nach absteigender Wichtigkeit gerankt. Da die einzelnen Fehler unterschiedliche Auswirkungen auf die Verständlichkeit

und damit die Brauchbarkeit einer Übersetzung haben, erscheint eine solche Gewichtung sinnvoll, um zu realistischen Aussagen über die Übersetzungsqualität zu gelangen. Dieses Ranking hat allerdings erst vorläufigen Charakter und ist an einer größeren Textauswahl oder auch am eigentlichen Testmaterial zu überprüfen.

Neben den oben genannten, zu berücksichtigenden Fehlertypen sind mit der Negation, Modalverben und Komposita bzw. Mehrwortausdrücken noch drei weitere Felder zu nennen, die bei fehlerhafter Übersetzung u.E. massive Auswirkungen auf das Satz- und Textverständnis haben können. Allerdings traten solche Fehler in den von uns untersuchten Texten nicht auf, so daß eine Entscheidung über eine Aufnahme dieser Kategorien in die Gruppe der zu berücksichtigenden Fehler noch aussteht, bzw. erst anhand der eigentlichen Fehleranalysetests gefällt werden kann.

Nicht zu berücksichtigende Fehlerklassen sind:

- Großschreibung:

Die Groß-/Kleinschreibung hat keinen Einfluß auf die Grammatikalität und damit die Verständlichkeit eines Satzes.

- Rechtschreibung

In der überwiegenden Mehrzahl der Fälle bleibt ein Wort in der Zielsprache auch bei fehlerhafter Rechtschreibung verständlich.

- Wortstellung (d.i. Satzgliedstellung)

Eine korrekte Segmentierung (s.o.) vorausgesetzt, bleibt ein Satz auch bei fehlerhafter Wortstellung verständlich.

- Auslassungen

Auslassungen treten praktisch nicht auf. Kann ein Wort oder eine Phrase nicht übersetzt werden, so wird die Originalform in die Übersetzung übernommen.

- Flexion (Kongruenz, Tempus)

Bei falscher Flexion bleibt ein Satz nach wie vor verständlich bzw. die Satzbedeutung rekonstruierbar.

- Passiv

Auch Fehler bei der Passivbildung bzw. eine fälschlicherweise nicht erfolgte Passivierung beeinträchtigen das Satzverständnis in der von uns untersuchten Textauswahl nicht massiv.

- Funktionswörter (auch Präpositionen und Konjunktionen)

Außer bei den lokalen Präpositionen kann die Bedeutung dieser Wortart aus dem Kontext erschlossen werden. Für die Konjunktionen scheint zu gelten, daß sich das logische Verhältnis zweier verknüpfter Sätze mehr aus dem semantischen Gehalt der Einzelsätze zueinander ergibt als aus der Bedeutung der Konjunktion selbst. Zudem traten in der von uns untersuchten Textauswahl keine Fehler bei der Übersetzung von Konjunktionen auf, so daß es vertretbar scheint, diese Fehlerart aufgrund mangelnder Relevanz unberücksichtigt zu lassen.

- Pronomen

Die Bedeutung der Pronomen ist meistens aus dem Kontext ersichtlich bzw. rekonstruierbar.

4.9 Evaluierungsziele: Produktivität

Die für die Produktivität eines automatischen Übersetzungssystems relevanten Faktoren können eingeteilt werden in die Gruppen Hardware, Software und anwenderspezifische Kriterien. Wie diese im Rahmen von MIROSLAV zu behandeln sind, soll im folgenden vorgestellt werden.

- **Hardware**

Hardwarefaktoren, die das Laufzeitverhalten eines Systems und somit seine Produktivität beeinflussen, sind der Prozessor, Taktfrequenz, benötigter Festplatten- und Arbeitsspeicher, Netzwerkfähigkeit sowie die Druckgeschwindigkeit des Druckers. Diese Daten sind anhand der Herstellerangaben leicht meßbar. Weiterhin entscheidend ist die Stabilität des Gesamtsystems, die im Rahmen der Benutzertests zur Produktivität der Software (s.u.).

- **Software**

Die Produktivität einer Software ist von einer Vielzahl von Einzelfaktoren abhängig, die jeweils einen mehr oder weniger großen Einfluß auf die Qualität und Quantität des Outputs nehmen. Folgende Tabelle, die sich bei der Einteilung an der ISO-Norm 9126 zu Qualitätscharakteristika von Softwareprodukten orientiert, gibt eine Aufstellung der im Rahmen der MIROSLAV-Evaluierung zu berücksichtigenden Kriterien.

Funktionalität	Verlässlichkeit	Benutzbarkeit	Effizienz	Portabilität
Texthandling (Import-/Exportmöglichkeiten) Einbindung in Arbeitsumgebung Netzwerkanbindung Übersetzungsarchive (Translation Memories) Lexikonhandling (Kodiertool) Einstellung von Sachgebieten Angabe von Übersetzungseinheiten Postediting-Unterstützung interaktives Übersetzen Voreinstellungen zur Übersetzung (z.B. grammatische Einstellungen wie Anrede)	Robustheit (Fehlertoleranz, Ausgereiftheit)	Mensch-Maschine-Interaktion (Bildschirmgestaltung, Verständlichkeit, Lernbarkeit, Operationalität) ⁶ Inkonsistenzvermeidung bei Lexikonediting Qualität der Systemdokumentation (Handbücher, Online-Hilfen etc.)	Geschwindigkeit	Installierbarkeit Plattformunabhängigkeit

⁶ Kriterien zur Softwareergonomie, wie sie bspw. in ISO 9126 genannt werden, sind aufgrund ihrer mangelnden Operationalisierbarkeit äußerst umstritten. Auf diese Diskussion kann allerdings im Rahmen der MIROSLAV-Evaluierung nicht eingegangen werden; es sei hier auf Krause/Womser-Hacker 1997 verwiesen. Das Spektrum der Mensch-Maschine-Interaktion wird bei der Produktivitätsanalyse von METAL anhand von Benutzertests untersucht.

Auf die Überprüfung der Wartbarkeit, zu der Kriterien wie Einfachheit der Fehleranalyse, Erweiterbarkeit u.ä. zählen, kann wegen der bereits erwähnten fehlenden diagnostischen Sichtweise der MIROSLAV-Evaluierung verzichtet werden.

Für die Vielzahl der softwareergonomischen Kriterien, die bei der Bewertung der Produktivität eine Rolle spielen, ist kein einheitliches Testdesign möglich. Hier sind je nach Art und Wichtigkeit des zu evaluierenden Bereichs unterschiedliche Verfahren anzuwenden wie z.B. direkte Evaluierung durch Auflistung der im System angebotenen Funktionen, analytische Bewertung der Benutzbarkeit durch eingehende Sichtung des Systems sowie Benutzertests mit Videoüberwachung und vorgegebenem Aufgabenset u.ä. (s. Marx 1998b).

- **Anwenderspezifische Kriterien**

Neben den Hardware- und Softwarekriterien gibt es noch eine Reihe systemexterner Faktoren, die die Produktivität eines maschinellen Übersetzungssystems beeinflussen. Dazu sind zu zählen:

Workflow

Wo und wie kann das Übersetzungssystem in den aktuellen Arbeitsablauf der Anwendungsdomäne eingesetzt werden? Sind durch den Einsatz eines maschinellen MT-Systems Produktivitätssteigerungen zu erwarten oder eingetreten? Hierbei sind nicht nur Quantitätssteigerungen beim Output ausschlaggebend, sondern auch Verbesserungen, die der Stabilität des gesamten Workflow, wie z.B. Dezentralisierungen dienen.

Textart

Die Produktivität eines Übersetzungssystems ist immer abhängig von der Art der zu übersetzenden Texte, da die einzelnen Textsorten von recht unterschiedlicher Komplexität sind. Daher ist die Angabe der bearbeitenden Textart unerlässlich. Da in MIROSLAV keine elaborierte Texttypologie geleistet werden kann, soll die Angabe einiger wesentlicher Eckdaten genügen:

- äußere Form: Ist das Dokument maschinenlesbar oder handelt es sich um einen Schreibmaschinentext?
- formale Kriterien: Wie hoch ist der Anteil an Graphiken, Tabellen, Aufzählungen, Fließtext etc.?
- grammatische Komplexität: Wie groß ist die durchschnittliche Satzlänge? Enthält der Text viele Einbettungen, komplexe Nominalphrasen, (Genitiv-)Attribute u.ä.?

Benutzermotivation

Die Motivation zur Nutzung eines maschinellen Übersetzungssystems beim MIROSLAV-Anwendungspartner kann nach Gesprächen vor Ort als sehr hoch eingestuft werden. Die mit der Bearbeitung russischer Texte betrauten Mitarbeiterinnen sind sehr am Einsatz einer solchen Komponente interessiert

und stehen ihr positiv gegenüber. Sie erwarten sich dadurch eine Arbeitsentlastung sowie eine Dezentralisierung bei der Arbeitsvergabe. Näheres ist durch einen Fragebogen zu ermitteln.

System- und Hardwarekenntnisse

Die System- und Hardwarekenntnisse beeinflussen die Produktivität eines Systems hauptsächlich insofern, als sie entscheidend sind für den Umgang der Benutzer mit Störsituationen. Sie sind durch einen Fragebogen zu ermitteln, in den auch Fragen zur Benutzermotivation eingehen sollen.

5 Zusammenfassende Übersicht über die MIROSLAV-Testverfahren

Nachfolgende Aufstellung gibt nochmals einen zusammenfassenden Überblick über die im Projekt MIROSLAV geplanten Evaluierungstests, getrennt nach Tests mit Benutzerpartizipation, sogenannte Benutzertests und denjenigen Testverfahren, die von MIROSLAV-Projektmitarbeitern durchzuführen sind.

Evaluation durch MIROSLAV-Projektmitarbeiter:

- Fehleranalyse
- Hardware- und Funktionalitätscheck
- Softwareergonomieanalyse
- Effizienztest
- Portabilitätstest

Benutzertests:

- Rating-Test
- Deskriptorentest
- Softwareergonomietest
- Workflowtest
- Umfrage zur Benutzermotivation

6 Ausblick

Mit dem Erscheinen der ersten PC-Version T1 für Deutsch-Russisch von Langenscheidt wird im Herbst diesen Jahres gerechnet. Sobald die Software vorliegt, kann mit der Testphase der MIROSLAV-Evaluierung begonnen werden. Bis Winter 1998/Frühjahr 1999 kann mit ersten Ergebnissen gerechnet werden, die zeigen, ob sich das theoretische Grundkonzept in der Praxis bewährt.

7 Literatur

[Bourbeau 1990]

Bourbeau, L. (1990): *Elaboration et mise au point d'une méthodologie d'évaluation linguistique de systèmes de traduction assistée par ordinateur*. Rapport final. Secrétariat d'État du Canada.

[Flanagan 1994]

Flanagan, M. (1994): *Error Classification for MT Evaluation*. AMTA 94.

[Hovy 1998]

Hovy, E. (1998): *Creating Useful Metrics for Evaluating Machine Translation*. Vortrag auf der First International Conference on Language Resources & Evaluation. Granada.

[Krause/Womser-Hacker 1997]

Krause, J.; Womser-Hacker, C. (1997): *Benutzerfreundlichkeit durch graphische Benutzungsoberflächen und die Integration von Vagheit*. In: Krause, J.; Womser-Hacker, C. (eds.): *Vages Information Retrieval und graphische Benutzungsoberflächen: Beispiel Werkstoffinformation*. Schriften zur Informationswissenschaft 28. Konstanz.

[Marx 1998a]

Marx, J. (1998a): *Vorarbeiten zur Evaluierung der Übersetzungsqualität und -produktivität von METAL (russisch-deutsch)*. In: MIROSLAV/R-Bericht 5/3. Regensburg.

[Marx 1998b]

Marx, J. (1998b): *Typologie der Evaluierung im Projekt MIROSLAV. Evaluierungskonzept, gewählte Testverfahren und -design*. MIROSLAV/R-Bericht 6/3. Regensburg.

[Marx/Mutschke/Schommler]

Marx, J.; Mutschke, P.; Schommler, M. (1995): *Möglichkeiten der intelligenten Integration heterogener Datenbestände - Das Projekt GESINE*; Dezember 1995; ISSN: 1431-6943; Herausgeber: Informationszentrum Sozialwissenschaften der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI).

<http://www.bonn.iz-soz.de/publications/series/working-papers/ab2/ab203.htm>

[Rinsche 1993]

Rinsche, A. (1993): *Evaluationsverfahren für maschinelle Übersetzungssysteme: zur Methodik und experimentellen Praxis*. Technical Report, Kommission der Europäischen Gemeinschaft, Bericht EUR 14766 DE.

[Rubio/Gallardo/Castro/Tejada 1998]

Rubio, A.; Gallardo, N.; Castro, R.; Tejada, A. (1998) (eds.): *First International Conference on Language Resources & Evaluation*. Proceedings. Granada.

[Van Slype 1979]

Van Slype, G. (1979): *Critical study of methods for evaluating the quality of machine translation*. Final report. Bruxelles: Bureau Marcel van Dijk.

[Vasconcello 1988]

Vasconcellos, M. (1988): *Factors in the Evaluation of MT: Formal vs. Functional Approaches*. In: Vasconcellos, M (ed.) (1988): *Technology as Translation Strategy*. Binghamton: State University of New York. pp. 203 - 213.

[Way 1994]

Way, A. (1994): *Developer-Oriented Evaluation of MT Systems*. In: Falkedal, K. (ed.): *Proceedings of the Evaluators' Forum, 1991, Les Rasses, Vaud, Switzerland*. Genf: ISSCO. pp. 237 - 244.

[White/Taylor 1998]

White, J.S.; Taylor, K.B. (1998): *A Task-Oriented Evaluation Metric for Machine Translation*. In: Rubio, A.; Gallardo, N.; Castro, R.; Tejada, A. (eds.): *First International Conference on Language Resources & Evaluation. Proceedings*. Granada.

Elektronische Wahlverfahren als Beispiel angewandter Kryptographie

Christian Wolff

Institut für Informatik
Universität Leipzig
Augustusplatz 10/11
D-04107 Leipzig
wolff@informatik.uni-leipzig.de

Inhalt

- 1 Einleitung
- 1.1 Plausibilität elektronischer Wahlverfahren
- 1.2 Vergleichbare Ansätze
- 2 Sicherheit und Kryptographie
- 3 Elektronische Wahlen
- 3.1 Ablauf des Wahlverfahrens
- 3.2 Systemaufbau
- 3.3 Implementierung
- 4 Sicherheitsbewertung
- 5 Generalisierung für beliebige Entscheidungsverfahren
- 6 Fazit
- 7 Literatur

Zusammenfassung

Der Beitrag schildert die Durchführung von Wahlen mit Hilfe kryptographischer Verfahren anhand einer Fallstudie mit Anwendungsbereich im Verbandswesen. Ausgehend von Vorüberlegungen zur kommunikativen Plausibilität und den technischen Voraussetzungen der Verwendung kryptographischer Algorithmen wird der Prototyp eines Wahlinformationssystems vorgestellt. Der Ansatz modelliert ein der traditionellen Briefwahl analoges Verfahren mit Hilfe von allgemein verfügbaren Kommunikationsdiensten im Internet (e-mail, WorldWide-Web). Dabei werden – im Vorgriff auf nach dem deutschen Signaturgesetz zertifizierte digitale Signaturen und Schlüssel – asymmetrische Schlüsselpaare nach dem RSA-Verfahren eingesetzt.

1 Einleitung

In diesem Aufsatz soll aufgezeigt werden, wie unter Einsatz angewandter Kryptographie Wahlverfahren elektronisch durchgeführt werden können. Die Untersuchung geht aus einer Pilotstudie für die *Gesellschaft für linguistische Datenverarbeitung* (GLDV) hervor, in der ein elektronisches Wahlverfahren realisiert und getestet wird (vgl. MIELKE & WOLFF 1998). In vielen Ländern werden derzeit rechtliche Rahmenbedingungen für den Einsatz kryptographischer Verfahren bzw. der digitalen Signatur geschaffen, so in Deutschland durch das Signaturgesetz (SigG) im Rahmen des Informations- und Kommunikationsdienstegesetzes (IuKDG) und die Signaturverordnung (SigV, vgl. BIESER 1997, BIESER & KERSTEN 1998:87-108). Die nachfolgend geschilderte Implementierung eines elektronischen Wahlverfahrens soll zumindest prinzipiell mit dem Einsatz kryptographischer Verfahren im Sinne des SigG kompatibel sein.

1.1 Plausibilität elektronischer Wahlverfahren

Wahlverfahren spielen in Vereinen und Verbänden zur Umsetzung von Entscheidungsprozessen eine wichtige Rolle und werden üblicherweise im Briefwahlverfahren durchgeführt, was mit nicht unerheblichen Kosten und einem hohen organisatorischen Aufwand verbunden ist. Die Motivation für den Einsatz elektronischer Wahlverfahren ergibt sich für derartige Personenmehrheiten u.a. aus folgenden Merkmalen:

Ein räumlich verstreuter heterogener Wählerkreis, der ein Briefwahlverfahren erforderlich macht.

Der Zugang zum Internet und seinen Diensten (e-mail, WWW) und die Realisierung bzw. Nutzung elektronischer Kommunikationsformen ist mit geringem Kostenaufwand möglich.

Traditionelle Kommunikationsverfahren bei Wahlen sind zeit- und kostenintensiv.

Die Wahlbeteiligung bei Verbandswahlen ist in der Regel relativ niedrig; ihre Erhöhung ist im Sinne einer höheren Mitgliederpartizipation anzustreben.

Diese Beobachtungen können als Argumente für eine Verwendung elektronischer Wahlverfahren gelten; zu beachten ist aber, daß auch mittelfristig nicht davon auszugehen ist, daß *alle* Wahlberechtigten an elektronischen Wahlverfahren teilhaben können oder wollen. Das nachfolgend beschriebene Verfahren hat daher *ergänzenden Charakter* und könnte allenfalls langfristig auf eine Ablösung traditioneller Kommunikationsformen abzielen. Rechtliche und demokratietheoretische Erwägungen zum Einsatz solcher Verfahren als neuer Form der Operationalisierung von Entscheidungsprozessen bleiben hier ausgeblendet (vgl. aber MIELKE & WOLFF 1998:32ff).

1.2 Vergleichbare Ansätze

Auch andere Fachgesellschaften bereiten derzeit elektronische Wahlverfahren vor oder haben sie bereits umgesetzt. Die *Association for Computing Machinery* (ACM) hat 1998 eine Satzungsänderung durchgeführt, die den Einsatz elektro-

nischer Wahlverfahren ermöglicht. Die *IEEE Computer Society* hat bereits ein elektronisches Wahlverfahren für ihre Verbandswahlen implementiert (<http://computer.org/election>). Die wichtigsten Merkmale dieses Verfahrens sind:

Verteilung eines Zugangsschlüssels (Kontrollnummer, *ballot control number*) auf den (gedruckten) Wahlunterlagen.

Realisierung der Wahlanwendung und der Durchführung der Wahl durch einen externen Dienstleister (ICE Interactive Certified Elections, <http://www.ice-ballot.com>).

Zugang zu einem WWW-Formular (Stimmzettel) mit Hilfe von Mitgliedsnummer und Kontrollnummer.

Versenden des Wahlzettels über das WWW (ohne Verschlüsselung).

Rückmeldung des Wahlzettels per e-mail (ohne Verschlüsselung).

Dieses Verfahren basiert auf einem kommerziellen Softwarepaket und setzt lediglich eine Zugangskontrolle zur Identifizierung des Wahlberechtigten ein; die Daten werden dabei mehrfach unverschlüsselt zwischen Wähler und Wahlurne bzw. Wahlamt hin- und hergesendet. Als Alternativen stehen die Wahl per Brief oder Fax zur Verfügung.

Auch für die elektronische Umsetzung politischer Wahlen entstehen derzeit erste Anwendungen. So wurde an der Universität Osnabrück ein Wahllokal für einen virtuellen Bundestagswahlkreis 329 eingerichtet, in dem man nach Zuteilung einer Wahlberechtigung eine Probeabstimmung zur Bundestagswahl 1998 durchführen kann. Der Ansatz versucht, das Briefwahlverfahren zu modellieren, verzichtet aber auf Verschlüsselungsverfahren (vgl. OTTEN 1998).

2 Sicherheit und Kryptographie

Für die Realisierung eines anonymen Wahlverfahrens mit elektronischen Mitteln spielen Aspekte der Sicherheit elektronischer Kommunikation eine zentrale Rolle. Dazu gehören u.a.

die Identifizierung (*authentication*) der Kommunikationspartner (Feststellung der Wahlberechtigung),

die Vertraulichkeit (*confidentiality*) der Kommunikation, insbesondere die Gewährleistung der Anonymität von Wahlen und

die Integrität (*integrity*) der kommunizierten Information (vgl. RANNENBERG, MÜLLER & PFITZMANN 1997:22f; SUN MICROSYSTEMS 1998).

Weitere Faktoren, die bei der Operationalisierung kryptographischer Verfahren Bedeutung haben, sind die *Transparenz* des Verfahrens und die *Vertrauenswürdigkeit* einer Anwendung. Die letzteren Kriterien dürften für die breite Durchsetzung kryptographiebasierter Kommunikation besonders wichtig sein, da nur durch sie hinreichende Benutzerakzeptanz zu erreichen ist. Zur Gewährleistung dieser Leistungsmerkmale dienen kryptographische Verfahren. Dabei ist zwischen unterschiedlichen Klassen kryptographischer Algorithmen (v.a. symmetrische und asymmetrische Kryptographie), ihrer softwaretechnischen Umsetzung in einem Dienst oder Protokoll (digitale Signatur, sicheres Kommunikati-

onsprotokoll, digitales Zertifikat, Datenverschlüsselung) und den angestrebten Schutzziele zu unterscheiden.

Das Grundprinzip der Kryptographie in ihrer klassischen Anwendung als Datenverschlüsselung ist dabei die Umwandlung einer zu übermittelnden Nachricht (*Klartext*, *plain text*, hier: der ausgefüllte Stimmzettel) mit Hilfe eines Kryptoalgorithmus in *Geheimtext* (*cyphertext*), der an den Empfänger gesendet und von ihm entschlüsselt werden kann. Neben *symmetrischen* Verfahren, bei denen Sender wie Empfänger über *denselben* Schlüssel zur Verschlüsselung bzw. Entzifferung verfügen, haben sich *asymmetrische* Verfahren durchgesetzt, bei denen der Sender den Klartext mit Hilfe des *öffentlichen* Schlüssels des Empfängers verschlüsselt. Der Empfänger kann die verschlüsselte Nachricht mit Hilfe seines *privaten* Schlüssels, der nur ihm bekannt und durch eine Paßwortphrase o.ä. geschützt ist, entziffern (Prinzip der *public key cryptography*, vgl. SALOMAA 1990:55ff, WOBST 1997:136ff). Auf der Basis *asymmetrischer* Verfahren lassen sich auch *digitale Signaturen*, das elektronische Pendant zur eigenhändigen Unterschrift, realisieren: Der Unterzeichner signiert ein Dokument mit Hilfe seines privaten Schlüssels. Unter Zuhilfenahme des öffentlichen Schlüssels kann der Empfänger verifizieren, daß ein Dokument von einem bestimmten Sender stammt und es *wie unterzeichnet* bei ihm eingetroffen ist (vgl. BIESER & KERSTEN 1998:20ff). Die Möglichkeit, einen Schlüssel zur Verschlüsselung von Daten problemlos veröffentlichen zu können, ohne die Sicherheit geschützter Daten zu kompromittieren, spielt auch bei dem hier erörterten Wahlsystem eine wesentliche Rolle.

3 Elektronische Wahlen

Das nachfolgend beschriebene Verfahren stellt die elektronische Umsetzung einer geheimen und anonymen Wahl auf der Basis einer Kombination symmetrischer und asymmetrischer Kryptographie dar. Dabei soll der Einsatz der kryptographischen Verfahren folgende Schutzziele gewährleisten:

Die *Anonymität* des Wählers, d.h. bei der Auszählung der Stimmzettel darf kein Hinweis auf die Identität des Wählers mehr vorhanden sein.

Die *Identifizierbarkeit* des Wählers, wofür eine digitale Signatur oder die Entschlüsselung mit Hilfe eines einem Wähler zugeordneten privaten Schlüssels zum Nachweis der Wahlberechtigung benötigt wird (Funktion des Wahlscheins bei der Briefwahl).

Die *Vertraulichkeit* nach außen, d.h. die ausgefüllten Stimmzettel müssen so versendet werden, daß kein Außenstehender von ihrem Inhalt Kenntnis erlangen kann und die

Sicherheit gegenüber Eingriffen unbefugter Dritter z.B. durch Schlüsselmißbrauch.

Die Realisierung eines solchen Verfahrens setzt voraus, daß

innerhalb der die Wahl durchführenden Vereinigung (Institution, Verein, Verband) die elektronische Wahlmöglichkeit durch geeignete Vereinbarungen (Satzungsänderung, Aufnahme der Möglichkeit einer elektronischen Wahl in eine Wahlordnung etc.) sanktioniert ist (rechtliche Voraussetzung),

die an der elektronischen Wahl teilnehmenden Wähler über e-Mail und WWW-Anschluß verfügen (technische Voraussetzung) und für alle anderen Wähler ein geeignetes „traditionelles“ Wahlverfahren zur Verfügung steht (Prinzip der parallelen Durchführung).

3.1 Ablauf des Wahlverfahrens

Aus einer Mitgliederdatenbank generiert das elektronische Wahlamt für jeden Wähler ein Schlüsselpaar. Die Schlüssel tragen einen impliziten Zeitstempel und sind nur für je eine Wahl gültig. Das Wahlamt fungiert dabei als Zertifizierungsstelle (vgl. BIESER 1997:402ff), d.h. es gewährleistet die Überprüfung von Schlüsseln und Signaturen auf ihre Gültigkeit. Bei der Verwendung von Schlüsselpaaren im Sinne des SigG würde diese Aufgabe von einem externen Dienstleister übernommen (*trust center*, vgl. BIESER & KERSTEN 1998:49ff), worauf hier aus Gründen der frühzeitigen Anwendbarkeit des Verfahrens verzichtet wird.

Der Wähler erhält seinen öffentlichen Schlüssel per e-mail zugesandt, während der private Schlüssel in der Datenbank des Wahlamts verbleibt. Zusätzlich generiert die Wahlsoftware auf der Clientseite mit einem symmetrischen Verfahren einen Schlüssel (Sitzungsschlüssel), mit dem nur der Wahlzettel verschlüsselt wird. Es kommt also in Analogie zur Briefwahl ein zweistufiges Verfahren zum Einsatz:

Der Wähler füllt den Wahlzettel aus, die Wahlsoftware verschlüsselt ihn mit Hilfe des symmetrischen Sitzungsschlüssels, ohne daß der Benutzer eingreifen müßte.

Der verschlüsselte Wahlzettel wird zusammen mit dem symmetrischen Sitzungsschlüssel nochmals mit dem wählerbezogenen öffentlichen Schlüssel verschlüsselt. Der öffentliche Schlüssel trägt damit die Funktion des legitimierenden Wahlscheins und dient der Identifizierung des Wählers.

Die Wahlsoftware sendet den so entstandenen Geheimtext an das Wahlamt. Es stellt zunächst anhand der verfügbaren privaten Schlüssel fest, ob der Wahlzettel von einem stimmberechtigten Wähler stammt. Anschließend legt es den verbleibenden, mit dem symmetrischen verschlüsselten Wahlzettel samt symmetrischem Schlüssel in die (elektronische) Wahlurne. Damit ist die Anonymität der Wahl bei gleichzeitiger Überprüfung der Wahlberechtigung gewährleistet. Nach Ablauf der Wahlfrist können alle eingegangenen anonymisierten und verschlüsselten Wahlzettel entschlüsselt und ausgezählt werden.

3.2 Systemaufbau

Aus der Schilderung des Wahlablaufs ergibt sich, daß zur Implementierung drei Softwarekomponenten erforderlich sind:

Das Modul für die Generierung und Verwaltung der Schlüsselpaare („Zertifizierungsstelle“) das auch die Identitätsüberprüfung und Anonymisierung eingehender Wahlbriefe übernimmt („Wahlamt“),

die eigentliche Wahlsoftware, ein WWW-Client, realisiert als Java-Applet („Wahlkabine“ mit „Stimmzettel“ (ein Java-Formular), der auch einen Container für den öffentlichen Schlüssel des Wählers enthält) und das Modul zur Entschlüsselung der Wahlscheine, das auch Auswertungsausgaben übernehmen kann („Wahlurne“).

Die konzeptuelle Trennung von „Wahlamt“ und „Wahlurne“ soll das Vertrauen der Wähler in das Verfahren erhöhen, da es den Aspekt der Anonymität der Wahl betont – ist der Stimmzettel in der Wahlurne, so ist kein Bezug zu einem bestimmten Wähler mehr vorhanden. Die Zuordnung zwischen den Verfahrensschritten und den Softwaremodulen sieht wie folgt aus:

Arbeitsschritt	Modul
Generieren wählerbezogener Schlüsselpaare	<i>Wahlamt</i>
Verteilen der öffentlichen Schlüssel per e-mail an die Wähler	<i>Wahlamt</i>
Ausfüllen, und symmetrisches wie asymmetrisches Verschlüsseln des Wahlzettels	<i>Wahlkabine</i>
Abschicken an den Wahlserver/Wahlamt	<i>Wahlkabine</i>
Verifikation der Wahlberechtigung (1. Entschlüsselung), Füllen der Wahlurne; Markieren des Schlüsselpaars als verbraucht	<i>Wahlamt</i>
Entschlüsseln der Stimmzettel (2. Entschlüsselung)	<i>Wahlurne/Auswertungsmodul</i>
Auswertung der Stimmzettel, Ergebnisaufbereitung	<i>Wahlurne/Auswertungsmodul</i>

Tab. 1: Zuordnung Arbeitsschritt - Modul

Die erste Entschlüsselung macht die Auswahl des passenden privaten Schlüssels aus dem Wählerverzeichnis erforderlich. Bei den zunächst erwarteten überschaubaren Teilnehmerzahlen ist dies durch systematisches Durchprobieren der Schlüssel möglich, wobei als Selektionsindikator ein nicht-informationstragendes Erkennungsmerkmal zur Eingrenzung des Schlüsselraums in der Datenbank mitgesendet werden kann. Bei späterer Verwendung von amtlich zertifizierten Schlüsseln entfällt das Problem, da an die Stelle der Datenbankabfrage zur Auswahl des passenden Schlüssels die Verifikation der digitalen Signatur des Wählers beim externen *trust center* tritt.

3.3 Implementierung

Grundsätzlich ließe sich diese Architektur mit Hilfe kryptographischer Standardsoftware realisieren. Dies hat aber organisatorische Nachteile, da dabei hohe Voraussetzungen an die beim Wähler vorhandene technische Infrastruktur gestellt werden (z.B. Installation von Softwarepaketen wie *Pretty Good Privacy* oder *Privacy Enhanced Mail - PEM*, vgl. IANNAMICO 1997; SCHNEIER 1996:664ff; WOOST 1997:275ff.) und Automatisierungs- wie Generalisierungspotentiale ungenutzt bleiben müssen. Um den Preis eines höheren Entwicklungsaufwands wird hier versucht, Benutzerfreundlichkeit durch ein einheitliches Interface mit wenigen Interaktionsschritten zu erreichen: Dem Benutzer steht für die Wahl der Stimmzettel als Java-Applet zur Verfügung, den er in einen Browser laden und ausfüllen kann. Der einzige unter ergonomischen Aspekten problematische

Verfahrensschritt ist die Übertragung des öffentlichen Schlüssels in den Wahlclient (Wechsel des Kommunikationsverfahrens von e-mail zum WWW).

Bei der Implementierung stehen folgende Aspekte im Mittelpunkt:

Für den im Umgang mit kryptographischen Verfahren unerfahrenen Benutzer steht ein einfaches Interface zur Verfügung, die Wahl muß mit wenigen Interaktionsschritten durchführbar sein (vgl. Abb. 1).

Da bei einem heterogenen Benutzerkreis keine klare Aussage über die beim Wähler vorhandene technische Infrastruktur zu treffen ist (Rechnerplattform, Browser), werden plattformübergreifende und weit verbreitete Dienste, Standards und Sprachen verwendet (e-mail, WWW/HTML, Java).

Eine weitgehende Prozeßautomatisierung soll den Arbeitsaufwand im Vergleich mit traditionellen Wahlverfahren minimieren (Schlüsselgenerierung und -verteilung, Entschlüsselung der Wahlscheine, Auswertung).

Sichere Kryptographie verlangt nach einem Höchstmaß an Transparenz. Deshalb kommen nur kryptographische Verfahren zum Einsatz, deren Algorithmen öffentlich einsehbar sind und für die verlässliche Einschätzungen ihrer Sicherheit existieren (RSA, IDEA).

Die eigentliche Wahlinformation wird auf der Basis einer XML-DTD im Klartext kodiert, so daß eine (auch gedruckte) Speicherung der anonymisierten Wahlscheine für die spätere Überprüfung möglich ist. Sie kann auch der Dokumentation der Wahlentscheidung für den Wähler dienen (vgl. Abb. 2).

Die modulare Struktur soll die Generalisierbarkeit der Anwendung sicherstellen.

Die Softwaremodule werden in Java implementiert (der WWW-Client als Java-Applet, die anderen Module als Java *Applications*), da so die Integration ins WWW erreicht werden kann. Zudem verfügt Java über ein generisches kryptographisches API, die *Java Cryptography Architecture* (JCA, vgl. LI GONG 1998). Dieses API modelliert auf einer abstrakten Ebene die benötigten kryptographischen Verfahren (Schlüsselerzeugung, Datenstrukturen für private und öffentliche Schlüssel, Implementierung der eigentlichen Verschlüsselungsalgorithmen). In dieses API kann die konkrete Implementierung der einzelnen Algorithmen (z.B. RSA, DSA, IDEA) eines *provider* eingehängt werden. Sun stellt zwar mit den *Java Cryptography Extensions* (JCE) einen *provider* mit den benötigten Algorithmen bereit; aufgrund der U.S.-amerikanischen Exportrestriktionen dürfen die JCE als eigentlich verschlüsselungsrelevanter Teil der JCA nicht aus den USA exportiert werden, da nach amerikanischen Recht Datenverschlüsselungsverfahren als *Waffen* gelten (vgl. SCHNEIER 1996:691ff, WOBST 1996:275ff, 307ff). Deshalb wird auf die Reimplementierung der Java JCA und der JCE der TU Graz zurückgegriffen (vgl. PLATZER 1998). Bei der Schlüsselerzeugung und der Auswertung kommt zudem das *Java Database Connectivity-API* (JDBC) zum Einsatz (Brücke zur Wählerdatenbank).

Die Klartextverschlüsselung des Stimmzettels verwendet den *International Data Encryption Standard* (IDEA), das derzeit wohl mächtigste kryptographische Verfahren (vgl. SCHNEIER 1996:370ff; WOBST 1997:182ff.). Die asymmetrische Verschlüsselung des Stimmzettels sowie der IDEA-Schlüssel werden mit dem öffentlichen Schlüssel des Wählers verschlüsselt. Für die Generierung der Schlüsselpaare des asymmetrischen Verfahrens verwendet der Client den

RSA-Algorithmus, das erste vollständige und am weitesten verbreitete asymmetrische kryptographische Verfahren (vgl. SALOMAA 1990:125-157, SCHNEIER 1996:531ff, WOBST 1997:143ff.). Ein Wechsel auf andere kryptographische Verfahren (z.B. von IDEA zu DES oder Triple-DES) ist aufgrund des hohen Abstraktionsgrads der in der JCA enthaltenen Klassen problemlos möglich – es muß nur jeweils ein Provider-Paket mit den benötigten Algorithmen zur Verfügung stehen.

Die Entscheidung, das Verfahren mit Hilfe asymmetrischer Kryptographie zu realisieren und Alternativen wie den Einsatz kryptographisch sicherer Übertragungsprotokolle (z.B. S/MIME, Secure Socket Layers oder S-HHTP, vgl. GARFINKEL & SPAFFORD 1997:120ff) zu vernachlässigen, beruht auf dem Ziel einer konzeptuellen Nähe zum zukünftigen Einsatz digitaler Signaturen im Sinne der deutschen Signaturgesetzgebung.

Abb. 1 zeigt den Prototyp des Webclients mit einem hypothetischen Wahlzettel. Er sieht Zustimmung, Ablehnung und Enthaltung als typische Wahlmöglichkeiten vor. Darüber hinaus kann für jede Einzelentscheidung auch *keine Stimmabgabe* erfolgen, wenn kein *radio button* selektiert ist.

4 Sicherheitsbewertung

Der vergleichsweise aufwendige Einsatz kryptographischer Verfahren ist nur berechtigt, wenn die heterogenen Sicherheitsanforderungen erfüllt werden können. Auf der Ebene der Algorithmen scheint nach derzeitigem Stand gewährleistet zu sein, daß nur ein *brute force*-Angriff auf den Geheimtext Erfolgchancen hat und die Menge unterschiedlicher Schlüssel nicht vollständig durchgerechnet werden kann.

Durch die genau begrenzte Menge verschiedener Schlüsselpaare, ihren eingeschränkten Einsatzzweck („Einmalschlüssel“) und ihre limitierte zeitliche Gültigkeit erscheint zudem ausgeschlossen, daß gefälschte Schlüssel in diesem Verfahren untergeschoben werden können. Gleiches gilt für das *Unterschieben* eines gefälschten Geheimtextes bei der Versendung des verschlüsselten Stimmzettels an das Wahlamt, das *Abfangen* und *Vernichten* eines Stimmzettels ist allerdings denkbar. Weitergehend wäre zu fragen, wie gut das Wählerverzeichnis, d.h. die Datenbank des Wahlamts mit den privaten Schlüsseln gegen Angriffe von außen geschützt ist.

PANEL

Ihr Stimmzettel zur Vorstandswahl

Name:

Vorname:

Ihr public key:

Name	Zustimmung	Ablehnung	Enthaltung
Livia, Anna	<input checked="" type="radio"/> Ja	<input type="radio"/> Nein	<input type="radio"/> Enthaltung
Myschkin, Lew	<input checked="" type="radio"/> <input type="text" value="Ja"/>	<input type="radio"/> Nein	<input type="radio"/> Enthaltung
Maultasch, Margarete	<input type="radio"/> Ja	<input checked="" type="radio"/> Nein	<input type="radio"/> Enthaltung
Biberkopf, Franz	<input checked="" type="radio"/> Ja	<input type="radio"/> Nein	<input type="radio"/> Enthaltung
Grandet, Eugenie	<input type="radio"/> Ja	<input type="radio"/> Nein	<input checked="" type="radio"/> Enthaltung

Verschlüsseln und Abschicken

Abb. 1: Interface der „Wahlkabine“

Ein gravierender Schwachpunkt ist offensichtlich die Versendung der öffentlichen Schlüssel per e-mail an die Wähler, da diese im Internet abgefangen und von Unbefugten zur Wahl verwendet werden könnten. Im Extremfall könnte bei Abfangen *aller* öffentlichen Schlüssel mit jedem abgefangenen Schlüssel genau ein Wahlzettel ausgefüllt werden.

Im Rahmen des gewählten Testszenarios erscheint dieser Fall hinreichend unwahrscheinlich, so daß auf weitere, das Verfahren verkomplizierende, Sicherungsmaßnahmen verzichtet wird. Mit diesem Argument der geringen Wahrscheinlichkeit eines Angriffs steht aber das Gesamtkonzept in Frage. Wichtiger erscheint daher die Perspektive des mittelfristigen Übergangs zu amtlich beglaubigten Schlüsselpaaren, der diese Schwachstelle eliminiert.

5 Generalisierung für beliebige Entscheidungsverfahren

Ein Abstimmungsverfahren wie oben geschildert setzt sich aus einer Menge unterschiedlich strukturierter Auswahlentscheidungen zusammen. Geht man davon aus, daß dieser *Verfahrensablauf* für unterschiedliche *Verfahrensinhalte* (Art und Anzahl der zu treffenden Entscheidungen) auf eine Vielzahl von Anwendungsfällen übertragbar ist, so wird offensichtlich, daß sich hinter der prototypischen Einzelanwendung ein großes Generalisierungspotential verbirgt. Mit Hilfe einer adäquaten Beschreibungssprache zur Spezifikation des Inhalts der Wahlzettel (oder allgemeiner: von Umfragen, Erhebungen etc.) wie einer XML-DDT (vgl. BRAY et al. 1997, LIGHT 1997) und einem geeigneten Parser läßt sich ein *Electronic Vote Management System* aufbauen, das nach Bedarf geeignete Wahlzettel generiert und im selben Format die Abstimmungsinhalte transportiert und auswertet. Dies ergibt einen weiteren Automatisierungsaspekt, nämlich die automatische Erstellung des Benutzerinterface für den Stimmzettel. Abb. 2 zeigt ein Beispiel eines XML-kodierten Stimmzettels, bei dem die Attribute des Elementes DECISION die Art der Entscheidung (NAME), ihren Typ (TYPE) und den Inhalt der Stimmabgabe (SELECTION) kodieren und automatisch generiert, geparkt und ausgewertet werden können. Für den Aufbau des Interfaces sind allerdings weitere Elemente und Attribute erforderlich, die den Wertebereich einer Einzelentscheidung definieren oder eine Binnengliederung bzw. -gruppierung der Einzelentscheidungen festlegen.

```
<VOTE>
  <DECISION NAME='LIVIA, ANNA'
    TYPE='SELECTION'
    SELECTION='YES' />
  <DECISION NAME='MYSCHKIN, LEW'
    TYPE='SELECTION'
    SELECTION='YES' />
  <DECISION NAME='MAULTASCH, MARGARETE'
    TYPE='SELECTION'
    SELECTION='NO' />
  <DECISION NAME='BIBERKOPF, FRANZ'
    TYPE='SELECTION'
    SELECTION='NO' />
  <DECISION NAME='GRANDET, EUGENIE'
    TYPE='SELECTION'
    SELECTION='ABSTENTION' />
</VOTE>
```

Abb. 2: XML-kodierter Stimmzettel

6 Fazit

Das dargestellte Verfahren wird derzeit als Prototyp entwickelt; ein erster Test soll im Herbst 1998 stattfinden, bevor 1999 der erste Einsatz für eine elektronische Vorstandswahl erfolgen kann. Dabei werden die Akzeptanz und die Nutzungsbandbreite wichtige Untersuchungsgegenstände sein, die auch Aufschluß

über die Plausibilität derartiger Verfahren für andere Einsatzzwecke geben können. Mittelfristig ist das Verfahren auf nach dem Signaturgesetz von einer amtlichen Zertifizierungsstelle beglaubigte digitale Signaturen umzustellen. Auf diese Variante wurde hier zunächst verzichtet, da davon auszugehen ist, daß noch sehr wenige Wahlberechtigte über eine solche Signatur verfügen.

7 Literatur

[ARNOLD & GOSLING 1997]

ARNOLD, Ken; GOSLING, James: *The Java™ Programming Language*, Reading/MA et al.: Addison-Wesley, 1997².

[BIESER 1997]

BIESER, Wendelin: *Begründung und Überlegung zum Signaturgesetz*. In: MÜLLER & PFITZMANN (1997), S. 399-410.

[BIESER & KERSTEN]

BIESER, Wendelin; KERSTEN, Heinrich: *Chipkarte statt Füllfederhalter. Daten beweissicher "elektronisch unterschreiben" und zuverlässig schützen*, Heidelberg: Hüthig 1998.

[BRAY 1997]

BRAY, Tim et al.: *Extensible Markup Language*. In: World Wide Web Journal 2(4) (1997), S. 29-68).

[GARFINKEL & SPAFFORD 1997]

GARFINKEL, Simson; SPAFFORD, Gene: *Cryptography and the Web*. In: World Wide Web Journal 2(4) (1997), S. 113-126.

[IANNAMICO 1997]

IANNAMICO, Mike: *Pretty Good Privacy™. PGP for Personal Privacy, Version 5.0 for Windows® 95, Windows NT. User's Guide*. San Mateo/CA: Pretty Good Privacy, Inc., 1997.

[KHARE & RIFKIN 1997]

KHARE, Rohit; RIFKIN, Adam: *Weaving a Web of Trust*. In: World Wide Web Journal 2(4), 1997, 77-112.

[LIGHT 1997]

LIGHT, Richard: *Presenting XML*, Indianapolis/IN: Sams.net, 1997.

[MIELKE & WOLFF 1988]

MIELKE, Bettina; WOLFF, Christian: *Kryptographiebasierte Kommunikationsformen für Vereine und Verbände*. In: LDV-Forum 15(1) (1988), S. 29-44.

[MÜLLER & PFITZMANN 1997]

MÜLLER, Günter; PFITZMANN, Andreas (edd.): *Mehrseitige Sicherheit in der Kommunikationstechnik. Verfahren, Komponenten, Integration*, Bonn et al.: Addison-Wesley, 1997.

[OTTEN 1998]

OTTEN, Dieter: *Aufruf zur Wahl im virtuellen Internet-Wahlkreis 329*. Universität Osnabrück, FB Sozialwissenschaften, <http://www.wahlkreis329.de/html/wahlauf.ruf.html>, August 1998.

[PLATZER 1998]

PLATZER, Wolfgang: *The IAIK Java Cryptography Extension*. TU Graz, Institute for Applied Information Processing and Communications, http://jcewww.iaik.tu-graz.ac.at/IAIK_JCE/jce.htm, Juni 1998.

[RANNENBERG, MÜLLER & PFITZMANN]

RANNENBERG, Kai; MÜLLER, Günter; PFITZMANN, Andreas: *Sicherheit, insbesondere mehrseitige IT-Sicherheit*. In: MÜLLER & PFITZMANN (1997), 21-29.

[SALOMAA 1990]

SALOMAA, Arto: *Public-Key Cryptography*, Berlin et al.: Springer, 1990.

[SCHNEIER 1996]

SCHNEIER, Bruce: *Angewandte Kryptographie*. Bonn et al.: Addison-Wesley, 1996.

[LI GONG 1998]

Li Gong (1998): *Java™ Security Architecture (JDK1.2)*. Draft Document (Revision 0.9). Palo Alto/CA: Sun Microsystems, <http://java.sun.com/products/jdk/1.2/docs/guide/security/spec/security-spec.doc.htm>, Juni 1998.

[SUN MICROSYSTEMS 1998]

SUN MICROSYSTEMS: *Secure Computing with Java: Now and the Future. A White Paper*. Palo Alto/CA: Sun Microsystems, <http://www.javasoft.com/marketing/collateral/security.html>, Juli 1998.

[WOBST 1997]

WOBST, Reinhard: *Abenteuer Kryptographie. Methoden, Risiken und Nutzen der Datenverschlüsselung*. Bonn et al.: Addison-Wesley, 1997.

Rhetorische Elemente in hypertextbasierten elektronischen Foren

Rolf Assfalg

Universität Konstanz
Fakultät für Mathematik und Informatik
Lehrstuhl Informationswissenschaft
78457 Konstanz
Rolf.Assfalg@uni-konstanz.de

Inhalt

1. Einführung
2. Das Modell eines argumentativen Internet-Hypertexts
3. Prototyp: Das KHS-Diskussionsforum
4. Rhetorische Elemente in KHS
5. Ausblick

Zusammenfassung

Zunächst wird der Prototyp eines elektronischen Diskussionsforums für das Internet beschrieben, wobei auf dessen Systemeigenschaften, die Implementation auf Basis des Konstanzer Hypertext-Systems (KHS) und auf bestehende Anwendungen eingegangen wird. Daran anschließend wird das Konzept einer Erweiterung dieses Systems beschrieben, mit dem es mehreren Benutzern zeitversetzt möglich ist, Beiträge abzusetzen, welche die rhetorischen Elemente einer Diskussion repräsentieren. Diese sind z.B. Thesen, Hypothesen, Synthesen, Argumente, Beispiele usf. Gleichzeitig mit dem Erzeugen eines solchen rhetorischen Elements ist der Benutzer gezwungen, die Art des Bezuges ausgehend von der aktuellen Position innerhalb der Diskussion mitanzugeben. Die Semantik der Bezüge ermöglicht dabei die Realisierung von Referenzen im Sinne von „Steht im Gegensatz zu“ oder „unterstützt“ usf. Das Neue an diesem Ansatz ist die Adaption der aus der Literatur bekannten rhetorischen Hypertext-Strukturen an die Erfordernisse, welche die Relevanz des Mediums Internet und seiner Bedeutung für die wissenschaftliche Kommunikation mit sich bringt.

1. Einführung

Mit der Einführung der Internet-Technologie wurde es durch den Dienst „Internet-News“ einem breiten Publikum möglich, an elektronisch geführten

Diskussionen teilzunehmen. Diesem Dienst kommt zwar eine hohe praktische Relevanz zu, jedoch werden zwischen den Teilnehmern nur Textfragmente ausgetauscht. Im Zuge längerer Diskussionen geht der Überblick verloren und der Benutzer ist gezwungen die einzelnen Beiträge zur lesen, um herauszufinden, ob ein Beitrag für ihn relevant ist, denn erst dann wird klar, ob der abgerufene Beitrag ein Problem beschreibt oder eine Lösung bietet. Verknüpfungen im Sinne von Hypertext-Verknüpfungen sind abgesehen von Frage-Antwort-Bezügen zwischen Beiträgen aber nicht möglich.

Neben Internet-News existieren jedoch einige aus der Hypertext-Literatur bekannte Systeme, mit denen sich Diskussionen mit hypertext-typischen Mitteln darstellen lassen. Eines der ersten argumentativen Hypertext-Systeme ist das System „glBIS“¹ von Conklin und Begeman (1988). Das System dient zur Aufzeichnung der Diskussionen einer Gruppe von Software-Entwicklern, und stellt die abgelegten Argumentationsstrukturen in farbig gestalteten Übersichtsgrafiken dar. Wenige Jahre später sind Systeme wie z.B. „Aquanet“ (vgl. Marshall et al. 1991), „SEPIA“ (vgl. Hannemann et al. 1992) und „HyperlBIS“ (vgl. Isenmann 1992) entstanden. „HyperlBIS“ dient zur Strukturierung von Information oder Streitfragen allgemeiner Art. Es ist als Ein-Benutzer-Anwendung implementiert. Bei SEPIA – einem weiteren anwendungsneutralen Hypertext-Prototyp - stand die Realisierung spezieller Autorenfunktionen zur Unterstützung kollaborativer Argumentation räumlich voneinander getrennter Autoren im Vordergrund. Auch bietet dieses System elaborierte Möglichkeiten zur Visualisierung des Argumentationsraums, der aus unterschiedlichen Perspektiven betrachtet werden kann, was den Benutzern durch Navigation entlang unterschiedlicher Pfade ermöglicht wird. Darüberhinaus stehen in „SEPIA“ eine Vielzahl von Bearbeitungs- und Übersichtswerkzeugen zur Verfügung. Die Gemeinsamkeit der genannten Systeme ist die Art der Unterscheidung zwischen Knoten und Verknüpfungen, wobei Knoten informationelle Einheiten repräsentieren, und Verknüpfungen einen semantischen Bezug zwischen diesen Knoten darstellen, der durch ihre Typzugehörigkeit – bei glBIS nur durch ein „Label“ - zum Ausdruck kommt. (s. Abbildung 1).

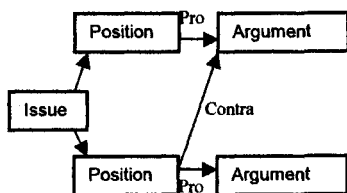


Abbildung 1: Hypertext-Argumentationsraum nach Conklin & Begeman (1988)

¹ „IBIS“ steht hierbei für „Issue-based Information System“. Ein von Kunz und Rittel (1979) entwickelter Ansatz.

Darüber hinaus sind einzelne Systeme von der klassischen Rhetorik und von Arbeiten aus der Argumentationstheorie inspiriert. So hat Toulmin (1975) im Zusammenhang mit der Erläuterung von Problemen des deduktiven Schließens eine graphische Notation zur Darstellung der inneren Struktur von Argumenten eingeführt (s. Abbildung 2), worauf bei der Entwicklung verschiedener Systeme Bezug genommen wurde.

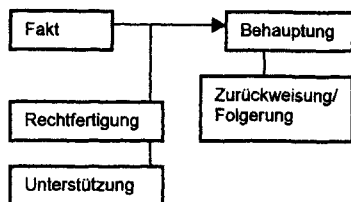


Abbildung 2: Argumentationsschema nach Toulmin (1975)

Eines der Systeme, bei denen auch Toulmin-Schemata berücksichtigt wurden, ist das System „Aquanet“, das am Xerox Palo Alto Research Center entwickelt wurde. Es dient unter anderem zur Strukturierung von Diskursen und von Wissen in ganz allgemeinem Sinne. Hierzu wird es den Benutzern ermöglicht, Typen von Informationsobjekten einzuführen, mit denen eine frame-artige Strukturierung der Knoten vorgegeben wird. Damit sind Typen zur Abbildung von hierarchischen Strukturen und nicht zuletzt auch von Toulmin-Schemata entstanden. Diese Toulmin-Argumente werden in „Aquanet“ als komplexe Relationen dargestellt, die technisch als Relationsobjekte realisiert sind, und die die einzelnen Bestandteile des Toulmin-Schemas (s. Abbildung 2) enthalten. Eine Unterscheidung zwischen Knoten und Verknüpfungen im klassischen Sinne gibt es bei „Aquanet“ nicht. Knotenobjekte enthalten hier nur Referenzen auf Instanzen von Basistypen, wohingegen Relationsobjekte Referenzen auf Instanzen aller Typen, also neben Instanzen von Basistypen, auch solche auf Knoten- und Relationsobjekte haben können. Verknüpfungen bei „Aquanet“ sind also n-stellig. Diese Vorgehensweise erscheint jedoch fragwürdig, da die Maxime der Modalität² von Verknüpfungen nicht gegeben ist, weil vom Benutzer beim Vorgang der Rezeption eine sehr komplexe Struktur auf einmal erfaßt werden muß. Mit diesem Modell werden also nur sehr erfahrene Benutzer zurechtkommen.

Toulmin-Schemata als Hypertext-Struktur nach Abbildung 1 darzustellen, wäre zwar prinzipiell denkbar, würde aber daran scheitern, daß Argumente bis auf Satzebene hypertext-artig zerlegt werden müßten. Aufwand und Nutzen stünden in keinem Verhältnis zueinander. Zudem würde die Übersicht leiden

² Die „Modalitätsmaxime“ von Verknüpfungen ist ein Begriff, der zum Ausdruck bringen soll, daß eine Verknüpfung ihre diskursiven Ziele klar zu erkennen geben sollte (vgl. Hammwöhler 96, S.54; Grice 93, S.249).

und es wären spezielle Mechanismen zur Gruppierung von Hypertext-Objekten notwendig.

Durch ihre eingeschränkte Portabilität und wegen ihres fehlenden Online-Zuganges sind die bisher erwähnten frühen Vertreter argumentativer Hypertexte („gIBIS“, „HyperIBIS“, „SEPIA“ und auch „Aquanet“) als Insellösungen einzustufen. In diesem Beitrag wollen wir eine Lösung skizzieren, die unter Verwendung des Konstanzer-Hypertext-Systems (KHS) – einem netzwerk-offenen Hypertext-System - die Möglichkeiten argumentativer Hypertexte im Internet präsent machen kann.

Ziel ist die Implementation eines Systems auf Basis des KHS, das die Vorteile der Netzwerk-Offenheit von Internet-News und die Idee der argumentativen Hypertexte zusammenbringen kann. Der Vorteil hierbei ist, daß mehrere räumlich getrennte Autoren sich nicht nur gegenseitig austauschen, sondern versuchen, gemeinsam Gedanken zu entwickeln oder Probleme zu erarbeiten.

Experimentelle Grundlage für das in den folgenden Abschnitten beschriebene System wird die Wissenschaftskommunikation sein, da hier ein Bedarf gesehen wird. Wissenschaftlicher Austausch findet einerseits in sehr formal statt. Viele halten es aber für mindestens genauso wichtig, auch informell zu kommunizieren, um sich fachlich auszutauschen. Allerdings - und das ist eine These - klafft eine „Lücke“, zwischen formaler und informeller Wissenschaftskommunikation. Es soll Publikation in kleineren Einheiten möglich sein, um die Hemmschwelle der Kommunikation geringer zu halten, um den Austausch von Information zu intensivieren. Daneben sollen synergetische Effekte der Kollaboration genutzt werden können. Motivation für das in diesem Beitrag beschriebene Vorhaben ist es, zur Schließung dieser Lücke etwas beizutragen. Möglicherweise wird damit auch die Grundlage für eine neue Generation elektronischer Zeitschriften geschaffen. Verbesserungen scheinen notwendig, wenn man berücksichtigt, daß elektronische Zeitschriften meist nur die Eins-Zu-Eins-Abbildungen ihrer gedruckten Vorbilder sind.

2. Ein Modell eines argumentativen Internet-Hypertext-Systems

Aus Benutzersicht soll die Präsentation der Inhalte aus o.g. Gründen eine Ähnlichkeit zum üblichen Seiten-Link-Paradigma erkennen lassen, um die Benutzer nicht zu überfordern³. Alternativ zur Verwendung n-stelliger Relationen, wird in der hier beschriebenen Lösung mit KHS vorgeschlagen, die in den Abbildungen 1 und 2 dargestellten Strukturen zu integrieren, indem man unter anderem toulmin-strukturierte Knoten einführt und diese als Ganzes mit anderen Toulmin-Argumenten, Positionen usw. verknüpft (siehe Abbildung 3). Wie im System „gIBIS“ von Conklin und Begeman (1988) (vgl. Abbildung 1 mit 3) stehen entsprechende Verknüpfungstypen zur Verfügung, um die semantischen Beziehungen der einzelnen Knoten zueinander zu definieren. Bei

³ Daher scheidet aufgrund der Nichterfüllung der „Modalitätsmaxime“ das Konzept der komplexen (n-stelligen) Verknüpfungsrelation aus, wie sie im System „Aquanet“ vorhanden ist.

der Eingabe stellen sich diese Knoten als Eingabe-Formular dar, so daß dem Benutzer intuitiv die Grenze zwischen Knoten und Verknüpfung klar wird.

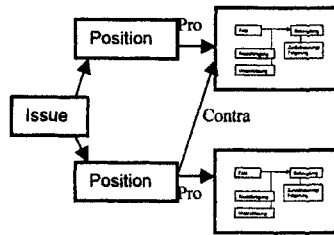


Abbildung 3: Argumentationsstruktur (nach Abbildung 2) mit strukturierten Knoten nach Toulmin (aus Abbildung 1).

Um den Benutzern verschiedene Zugangspfade zu ermöglichen, verfügt das Hypertext-Modell des KHS über Möglichkeiten der polyhierarchischen Strukturierung von Knoten durch spezielle Hierarchie-Verknüpfungen. Diese Möglichkeiten der Strukturierung können zur Organisation von Argumentationsstrukturen genutzt werden. Insbesondere die Polyhierarchie ermöglicht die gleichzeitige Einbettung von Argumentationsstrukturen in verschiedene thematische Kontexte.

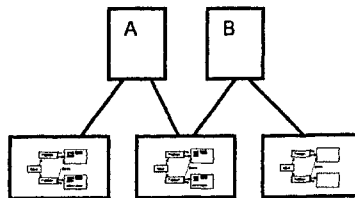


Abbildung 4: Polyhierarchische Strukturierung von Argumentationsräumen in KHS

Im folgenden wollen wir zunächst eine prototypische Implementation eines diskursbereichsübergreifenden Diskussionsforums betrachten, das auf der Basis des Konstanzer Hypertext-Systems (KHS) entstanden ist. Danach wird auf die geplanten Erweiterungen des Systems in Richtung argumentativer Hypertexte eingegangen.

3. Prototyp: Das KHS-Diskussionsforum

Mit der Prototyp-Implementation eines Diskussionsforums auf der Basis des Konstanzer Hypertext-Systems (KHS) (vgl. Hammwöhner 1996 und Rittberger et al. 1994, Assfalg et al. 1997) wird derzeit auf zwei elektronischen Marktplätzen experimentiert. Einzelne Foren sind dem Internet-Publikum über die Internet-Schnittstelle des KHS zugänglich⁴. Abbildung 5 vermittelt einen Eindruck, wie sich das Forum für den Internet-Benutzer darstellt. KHS ist in Smalltalk implementiert und besitzt als Datenspeicherungskomponente die objektorientierte Datenbank GemStone, deren Programmier- und Abfragesprache ebenfalls Smalltalk ist. Der Internet-Zugang zu KHS-Inhalten wurde über die Implementation eines HTTP-Listeners auf der GemStone-Datenbank realisiert.

Die Diskussionsfunktionen werden KHS-intern über spezifische Knoten- und Verknüpfungstypen realisiert. Ein einzelnes Forum wird zunächst durch einen aggregierten KHS-Knoten abgebildet, welcher Diskussions-Rubriken enthält. Diese wiederum werden durch Instanzen eines speziellen aggregierten KHS-Knoten-Typs gebildet. In den Rubriken sind dann die Beiträge enthalten. Im derzeitigen Entwicklungsstadium des Systems können dem Benutzer „Hauptbeiträge“, „Anonyme Beiträge“ und „Abstimmungen“ angeboten werden. Daneben gibt es noch ein Verzeichnis der in diesem Forum registrierten Benutzer, die genauso wie die Diskussionsbeiträge als Hypertext-Knoten repräsentiert werden. Benutzerknoten speichern Benutzername, E-Mail-Adresse und Home-Page sowie Daten zur Kontrolle des Systemzugangs. Eine Mail-Benachrichtigungsfunktion teilt den registrierten Benutzern mit, wenn neue Beiträge eingegangen sind.

⁴ Eine Referenzanwendung ist: <http://khs.uni-konstanz.de:57786/getKHS/Alpha1/emb> Hierbei handelt es sich um eine „Bürgerzeitung“ für die elektronischen Marktplätze „Electronic Mall Bodensee (EMB)“, <http://www.emb.net> und „Bodensee 24“ <http://www.bodensee24.de>

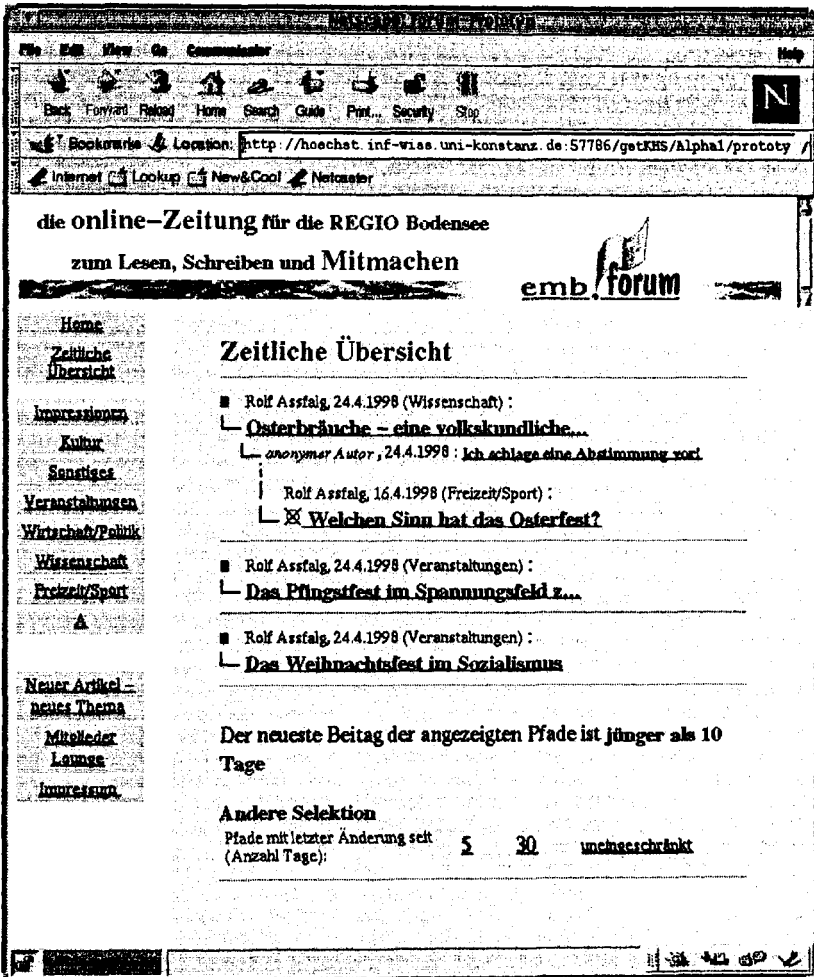


Abbildung 5: KHS-Diskussionsforum in Form einer Bürgerzeitung für die Elektronik Mall Bodensee (EMB) – einem regionalen, elektronischen Marktplatz. Zu sehen sind Standardbeiträge, ein anonymer Beitrag und eine Abstimmung.

Für redaktionelle Tätigkeiten stehen unter KHS Editierwerkzeuge zur Verfügung. Diese ermöglichen die Manipulation von Inhalten und Meta-Daten. Es können z.B. einzelne Beiträge per „Drag and Drop“ gleichzeitig mehreren Verzeichnissen zugeordnet werden. Auf diese Weise ist ein Knoten über mehrere Pfade erreichbar. Eine WWW-Schnittstelle ermöglicht den Zugriff auf die Inhalte vom Internet her.

Die einzelnen Dokumente in KHS-Diskussionen sind über Hyperlinks spezifischen Typs untereinander verknüpft. So können beispielsweise Beiträge mit anderen Beiträgen annotiert werden. Die Annotationsbeziehung zwischen den beiden Beiträgen wird dabei vom Hypertext-System über eine Verknüpfung des Typs „IsAnnotatedBy“ repräsentiert. Da auch Kommentare wieder kommentiert werden können, ist der Aufbau von Hierarchien aus thematischen Bezügen möglich. Eine solche Hierarchie, die ein Thema widerspiegelt, nennen wir „Thread“ (s. Abbildung 5)

Thread-Übersichten zeigen das Verknüpfungsgeflecht, das sich im Zuge der Entstehung einer sich hierarchisch ausbildenden Kommentierungsstruktur ergibt. Solche Übersichten werden zusammen mit der Darstellung der einzelnen Beiträge angezeigt und sind technisch mit HTML-Tabellenkonstrukten realisiert. Sie können sich jedoch im Zuge lebhaft geführter Diskussionen sehr stark aufblähen. Aus diesem Grund wurde ein Fish-Eye-ähnlicher (vgl. Furnas 1996) Filter implementiert, der die Thread-Übersicht auf ca. 10 Einträge begrenzt. Für jeden Beitrag wird dazu anhand bestimmter Kriterien eine Gewichtung errechnet.

Offene Hypertext-Systeme wie KHS sind als Basis einer technisch erweiterbaren Diskussionsplattform geeignet, da die unterschiedlichen Arten von Beiträgen durch typisierte Knoten abgebildet werden können. Im folgenden wollen wir das Konzept einer Erweiterung der Typ-Menge von KHS-Knoten und Verknüpfung beschreiben, mit der es möglich ist, rhetorische Strukturen zu beschreiben.

4. Rhetorische Elemente in KHS

Bei Betrachtung der KHS-Prototyp-Implementation erscheint die Unterscheidung einzelner Diskussionsbeiträge nach anonymen Beiträgen, Hauptbeiträgen und Abstimmungen noch etwas grob. Daher sind Erweiterungen geplant, die es erlauben, mittels einer feineren Gliederung der Beitragstypen eine Diskussion zu führen, deren innere Bezüge über die Bildung von Kommentierungshierarchien hinausgehen. Damit entstehen netzwerkartige Diskussionsverläufe, was der Hypertext-Idee entgegenkommt.

Konkret geplant ist es, aus dem Beitragstyp „Hauptbeitrag“ zunächst folgende zusätzliche Sub-Typen zu bilden:

Diskussionsanstoß

Position/These

Synthese

Hypothese

Argument (Text)

Toulmin-Argument

Faktum/Beleg/Zitat

Beispiel

Anmerkung/Ergänzung/Frage/Aufforderung

Die Eingabe der Inhalte in Dokumente der o.g. Typen wird mit Formularen realisiert. Hierbei läßt sich die Eingabe der Benutzer steuern und es kann durch entsprechende Beschriftung der Eingabefelder auch ein Toulmin-Schema (z.B. für Knotentyp „Argument“) nach Abbildung 1 vorgegeben werden.

Zur Realisierung rhetorischer Bezüge zwischen den Ausprägungen dieser Beitragstypen von Beiträgen sind folgende Verknüpfungstypen geplant:

steht im Gegensatz zu

unterstützt (Argument/These)

illustriert/macht plausibel

führt zu These

proviziert (Anmerkung/Ergänzung/Frage/Aufforderung)

Im Verlauf der Diskussion steht dem Teilnehmer zum Erzeugen eines neuen Beitrags immer nur eine Teilmenge der gesamten Beitragstypen zur Verfügung, was der Bedienungsfreundlichkeit des Systems entgegenkommt. Zur Auswahl eines oder mehrerer innerhalb eines konkreten Navigationskontexts zur Verfügung stehenden Typen gelten folgende Regeln:

Eine Diskussion kann mit dem Beitragstyp „Diskussionsanstoß“ oder durch eine Verknüpfung, die von einem beliebigen Fachbeitrag ausgeht, begonnen werden.

Auf einen Fachbeitrag, „Diskussionsanstoß“ oder auf ein Beispiel kann nur in Form von Hypothesen, Positionen/Thesen, Fakten und Anmerkungen Bezug genommen werden.

Eine Synthese kann sich nur auf zwei Thesen beziehen.

Anmerkungen können sich nur auf andere Anmerkungen, Argumente und Beispiele beziehen.

Eine Abstimmung kann sich nur auf Thesen oder Synthesen beziehen. Eine Abstimmung soll den dramaturgischen Schlußpunkt einer Diskussion setzen.

Auch die Verknüpfungsregeln sind vom Benutzer zu wählen. Natürlich gilt auch hier wie bei der Auswahl von der Beitragstypen, daß immer nur eine Teilmenge der im System vorhandenen Verknüpfungstypen angeboten wird. Zur Einschränkung der Menge der dem Benutzer im Dialog zur Auswahl angebotenen Verknüpfungstypen gelten die folgenden Regeln:

Die „Im-Gegensatz-Zu-Beziehung“ kann nur zwischen Argumenten und Positionen installiert werden.

Die „Unterstützung“ kann nur auf Belege und Argumente hinführen.

Die „Führt-Zu-These-Beziehung“ zielt nur auf Positionen oder Thesen.

die „Illustriert“- oder „Macht-Plausibel-Beziehung“ führt zu Beispielen oder Belegen.

Die „Proviziert-Beziehung“ kann von Beiträgen jeglichen Typs ausgehen und zeigt auf Anmerkungen, Ergänzungen, Fragen oder Aufforderungen.

Dieses Regelwerk, das dem Benutzer in einer konkreten Situation immer nur eine Teilmenge der im System vorhandenen Knoten- und Verknüpfungstypen zur Auswahl gibt, bringt kognitive Entlastung und führt gleichzeitig zu einem in sich plausiblen Gebrauch der rhetorischen Stilmittel. Das Fehlen einer expliziten Kommentarverknüpfung ist aufgrund der Erfahrungen von Conklin und Begeman (1988) notwendig, da im Zuge des experimentellen Gebrauchs des Systems „gIBIB“ festgestellt wurde, daß die Benutzer eher dazu neigen, die Beiträge anderer Teilnehmer zu kommentieren oder zu ergänzen, anstatt Positionen, Gegenpositionen oder Argumente zu formulieren.

5. Ausblick

Eine besondere Herausforderung stellt die Visualisierung von Thread-Hierarchien in argumentativen Hypertexten dar. Auf den ersten Blick erscheint es sinnvoll, analog zur Visualisierung der Thread-Hierarchien vorzugehen, wie sie in Abbildung 5 dargestellt ist. Jedoch verschließt man sich hierbei der Möglichkeit, Diskussionen in Form eines gerichteten azyklischen Graphen (DAG) zu erzeugen. Damit kann ein einzelner Knoten über mehrere Pfade erreicht werden. Er kann damit sozusagen in mehreren thematischen Kontexten stehen. Solche komplexen Graphen möglichst kreuzungsfrei und in einer Art und Weise zu präsentieren, daß der Benutzer in der Lage ist, das Wesentliche möglichst rasch zu erfassen, berührt Fragestellungen aus der Graphentheorie und der Kognitionspsychologie.

Zur Implementation der Visualisierungsfunktionalität eines argumentativen Hypertexts für das Internet reichen die Ausdrucksmöglichkeiten von HTML nicht aus. Es ist damit notwendig, solche Visualisierungskomponenten als JAVA-Applets zu realisieren. Durch deren Anwendung sollte es den Benutzern auch möglich sein, die graphische Präsentation der vorliegenden Argumentationsräume zu beeinflussen. Nach Marshall et al. (1991) ist dies einerseits eine wichtige Voraussetzung, um eine gute visuelle Adaptivität zu erhalten. Sie bringt andererseits jedoch die Gefahr mit sich, daß Benutzer dazu verleitet werden, Beziehungen zwischen Hypertext-Knoten nur durch räumliche Anordnung und nicht durch das Verknüpfen derselben auszudrücken. Eine Lösung für einen argumentativen KHS-Hypertext für das Internet könnte sein, die räumliche Anordnung von Knoten ausschließlich einem Moderator zu überlassen. Eine andere Möglichkeit wäre es, Verknüpfungen einzuführen, welche die räumliche Nähe zweier Knoten-Objekte auf der Visualisierungskomponente repräsentieren, die jedoch nur system-intern existieren jedoch auf der Oberfläche der Visualisierungskomponente nicht dargestellt werden.

Ermöglicht man Benutzern das Einführen von Knoten- und Verknüpfungstypen, wie es im System „Aquanet“ der Fall ist, entstehen nach (Marshall et al. 1991) sehr unterschiedlich gestaltete Schemata, was sehr stark vom diskutierten Gegenstand und von den an der Diskussion beteiligten Personen abhängt. Auch hier scheint eine Beschränkung der Funktionen zur Wartung der

Verknüpfungstypen und der Knotenschemata auf Benutzer mit Moderatöraufgaben sinnvoll.

Zu erkennen, daß z.B. zwei Argumente in Gegensatz zueinander stehen, kann für sich genommen bereits als Leistung eines Autors angesehen werden. Dieses ohne einen Textbeitrag zu erstellen, also nur durch das bloße Erzeugen einer Verknüpfung zu dokumentieren, welche diese Gegensätzlichkeitsbeziehung ausdrückt, ist konsequenterweise ein wissenschaftlicher Beitrag. Mit Systemen, in denen solche Möglichkeiten bestehen, wandelt sich der Autorenbegriff und berücksichtigt auch Handlungen, die man momentan eher noch mit Begriffen wie „Editieren“ oder „Verlinken“ umschreiben würde. Dies wirft natürlich Fragen der Bewertung der von den Autoren erbrachten Leistungen und Nutzungsrechte auf. Es ist damit auch eine Herausforderung an die Schreibforschung.

Aufgrund der höheren Verfügbarkeit von JAVA-Laufzeitumgebungen, der größeren Verbreitung und Attraktivität dieser Sprache und aufgrund der Tatsache, daß diese Sprache für Internet-Anwendungen geradezu prädestiniert ist, wird das bislang in Smalltalk-Umgebungen laufende Konstanzer Hypertext-System (KHS) derzeit in JAVA reimplementiert. Damit ist die Grundlage geschaffen, um die Entwicklung der offenen Hypertext-Systeme in ihrer speziellen Ausprägung als Internet-Anwendungen voran zu bringen.

Literatur

[Assfalg et al. 97]

Assfalg, R./ Goebels, U./ Welter, H. (1997): *Internet-Datenbanken – Konzepte, Methoden, Werkzeuge*. Addison Wesley Longman. Bonn, 1997

[Conklin & Begeman 88]

Conklin, J., and Begeman, M.L. (1988): *gIBIS: A Hypertext Tool for Exploratory Policy Discussion* in: ACM transactions office information systems 6,4 (Oktober 1988), 303-331. Auch in: Proc. 2nd conference computer supported cooperative work (Portland, OR, 26.-28. September 1988), S.140-152.

[Furnas 86]

Furnas, G.W. (1986): *Generalized fisheye views* in: Proceedings of the ACM CHI '86 Conference on Human Factors in Computing Systems. Boston (MA), 27-42

[Grice 93]

Grice, H.P. (1993): *Logik und Konversation* in: Meggle, G.: *Handlung, Kommunikation, Bedeutung*, S.243-265, Suhrkamp, 1993

[Hammwöhner 96]

Hammwöhner, R. (1996): Offene Hypertextsysteme: Das Konstanzer Hypertext-System im technischen und wissenschaftlichen Kontext. Habilitationsschrift. Universität Konstanz, Informationswissenschaft, 1996

[Hammwöhner & Kuhlen 94]

Hammwöhner, R./ Kuhlen, R. (1994): *Semantic control of open hypertext systems by typed objects*. In: Journal of Information Science, Vol. 20, Nr. 3: 175-184, 1994

[Hannemann et al 92]

Hannemann, J./Thüring M./Friedrich N. (1992): *Hyperdocuments as user interfaces: Exploring a browsing semantic for coherent hyperdocuments* in: Cordes, R./Streitz, N. (Hg.) Proceedings „Hypertext und Hypermedia 1992 - Konzepte und Anwendungen auf dem Weg in die Praxis“, Springer-Verlag, Heidelberg, 87-102, 1992

[Isenmann 92]

Isenmann, S. (1992): *HyperIBIS - a Tool for Argumentative Problem Solving*. In: A. Min Tjoa, Isidro Ramos (Eds.): Proceedings of the International Conference „Database and Expert Systems Applications“ in Valencia, Spain. Springer-Verlag, Wien, 1992

[Kunz & Rittel 79]

Kunz, W./ Rittel, H. (1979): *Issues as Elements of Information Systems* (Working Paper 131), Center for Planning and Development Research. University of California, Berkeley, 1979

[Marshall et al 91]

Marshall, C.C. / Halasz, F.G. / Rogers, R.A. / Jannssen, W.C. Jr. (1991): *Aquanet: a hypertext tool to hold your knowledge in place*. Proceedings of ACM Hypertext 91, Dec. 1991 San Antonio, S. 261-275, ACM New York, 1991

[Rittberger et al. 94]

Rittberger, M./ Hammwöhner, R./ Assfalg, R./ Kuhlen, R. (1994): *A homogenous interaction platform for navigation and search in and from open hypertext systems*. Conference Proceedings RIAO '94: Intelligent Multimedia Information Retrieval Systems and Management. New York Okt. 1994, S. 649-663

[Toulmin 75]

Toulmin, S. (1975): *Der Gebrauch von Argumenten*, Scriptor Verlag, Kronberg/Ts. Deutsche Übersetzung der englischen Ausgabe: *The Uses of Argument*, Cambridge University Press, 4. Auflage (Erstausgabe. 1958)

Selix im DFG-Projekt Kascade

Hubert Hüther

SOFTEX GmbH
Hubert Hüther
Schmollerstr. 31
66111 Saarbrücken
Email: hhuether@softex.de

Inhalt

1. Einleitung
2. Der Rahmen von SELIX
 - 2.1 KASCADE
 - 2.2 MILOS
3. Die Aufgabe von SELIX
4. Die Gewichtung nG bei SELIX
 - 4.1 Das Kollektionsgewicht nG1
 - 4.2 Das Dokumentgewicht nG2
 - 4.3 Das Längengewicht nG3
5. Das Programmsystem von SELIX
6. Ausblick
7. Literatur

1 Einleitung

Mit KASCADE wird der Aufbau einer Datenbasis aus inhaltlich angereicherten und maschinell tiefer gehend erschlossenen Dokumenten (über Titel und bibliographische Daten hinaus) durchgeführt.

Mit dem Indexierungssystem MILOS/IDX (LZ 97) werden die Daten zunächst erschlossen.

Mit dem System SELIX werden, ausgehend von diesen Ergebnissen von MILOS, Parameter für einzelne Dokumente und auch die gesamte Kollektion der Dokumente ermittelt. Mit Hilfe dieser dann vorhandenen Parameter werden weitgehend automatisch Entscheidungen gefällt, ob ein von MILOS erschlossener Indexterm für ein Dokument zur Indexierung benutzt werden soll. Gleichzeitig steht dem Benutzer ein breites Spektrum zur Steuerung zur Verfügung.

2 Der Rahmen von SELIX

2.1 KASCADE

Seit dem 1. April 1997 wird mit Förderung der Deutschen Forschungsgemeinschaft an der Universitäts- und Landesbibliothek Düsseldorf das Projekt KASCADE durchgeführt. Gemeinsam mit der Fachrichtung Informationswissenschaft der Universität des Saarlandes sollen in dem auf eine Laufzeit von 21 Monaten (ursprünglich 18 Monate) angelegten Projekt neue Möglichkeiten der Inhaltserschließung und des Information Retrieval entwickelt und im bibliographischen Einsatz getestet werden.

Ziel von KASCADE ist der Aufbau einer Datenbasis aus inhaltlich angereicherten und maschinell tiefer gehend erschlossenen Dokumenten (über Titel und bibliographische Daten hinaus). Die Welt des klassischen bibliothekarischen Online-Katalogs (OPAC) wird dabei also verlassen.

Grundlage für die Datenbasis von KASCADE und den entsprechenden Demonstrator ist das Fachgebiet Jura (Zeitschriften und Monographien) aus dem Datenbestand der Universitäts- und Landesbibliothek Düsseldorf, das hinsichtlich Größe (ca. 30.000 Titel) und sprachlicher Zusammensetzung (großer deutschsprachiger Anteil) für die Zielsetzung gut geeignet ist.

Für die Anreicherung der Erschließungsdaten wurden für etwa 3.000 Dokumente die Inhaltsverzeichnisse gewählt (da diese im Unterschied zu Abstracts und Registern fast immer vorhanden waren) (siehe auch (LZ 97)).

2.2 MILOS

Mit dem Indexierungssystem MILOS/IDX (LZ 97) werden die Daten zunächst erschlossen. Das Eingabeformat für dieses Indexierungssystem ist das ASCII-Format. Dabei können Zusatzinformationen in Form von Kommentaren mitgegeben werden. Für die Auswertung der Indexierungsergebnisse (ebenfalls ASCII-Format) ist für SELIX die Dokumentnummer als Zusatzinformation erforderlich. Bei der Indexierung werden Grundform, Wortklasse, Teilwörter usw. (siehe (LZ 97)) gewonnen (dabei steht ein Relationenwörterbuch (Thesaurus) zur Verfügung). Die dann vorhandenen Indexierungselemente für ein Dokument können sehr zahlreich sein.

3 Die Aufgabe von SELIX

Mit dem System SELIX (Selektive automatische Indexierung) werden, ausgehend von den Ergebnissen von MILOS unter Benutzung eines Relationenwörterbuches, Häufigkeiten für Grundformen und Parameter für einzelne Dokumente und auch die gesamte Kollektion der Dokumente ermittelt (dabei können unter Benutzung des Relationenwörterbuches (Thesaurus) auch verschiedene Grundformen mit Hilfe von Synonym-, Ähnlichkeits-, Teilwort-, Vorzugsrelationen und/oder weiterer Relationen zusammengeführt werden). Mit Hilfe dieser dann vorhandenen Parameter sollen weitgehend automatisch Entscheidungen gefällt werden, ob ein von MILOS erschlossener Indexterm für ein Dokument zur Indexierung benutzt werden soll. Hierzu wird für jede in einem

Dokument (vorkommende oder zugeteilte Grundform) ein auf dieses Dokument bezogenes Gewicht ermittelt.

Im Unterschied zu aus der Literatur bekannten Gewichtsfunktionen (z.B. in HA 75, SA 89 oder HR 98) sollen auch bestimmte selten vorkommende Grundformen (komplexe Mehrwortgruppen oder Komposita) ‚statistisch gesehen‘ eine Chance erhalten. Ebenso soll der Benutzer sogenannte ‚Registerbegriffe‘ festlegen können (diese Begriffe werden einem Dokument zugeteilt, wenn sie im Dokument vorkommen oder mit einem im Dokument vorkommenden Begriff durch eine entsprechende Relation verbunden sind (unabhängig von seinem ‚Gewicht‘)). Das statistische Verfahren muss die Massenverarbeitung erledigen, aber der Benutzer soll ein umfangreiches Spektrum (Parameter, Relationenwörterbuch (Thesaurus)) zur Steuerung zur Verfügung haben.

Von SELIX werden hierzu drei Teilgewichte ($nG1 - nG3$) ermittelt. Die drei Teilgewichte sollen eine Maßzahl zu einem Begriff liefern:

- ($nG1$) Eignung als Indexterm für die Dokumentkollektion (kollektionsbezogen)
- ($nG2$) Wichtigkeit des Begriffes für ein Dokument (dokumentbezogen)
- ($nG3$) Bedeutung des Terms für das Fachgebiet.

Die drei Gewichte ergeben additiv das Gesamtgewicht, wobei jedes Gewicht noch mit einem Faktor versehen werden kann.

4 Die Gewichtung nG von SELIX

Folgende Parameter werden mit SELIX ermittelt:

- $nCollen$: Länge der Dokumentenkollektion bezogen auf die erfassten Wortformen.
- $nAnzDok$: Anzahl der Dokumente in der Kollektion.
- $HfklmD(g,d)$: Häufigkeit einer Grundform 'g' in einem Dokument 'd'.
- $nDokLen(d)$: Dokumentlänge von 'd' bezogen auf die erfassten Wortformen.
- $nAnzGru(d)$: Anzahl der erfassten verschiedenen Grundformen im Dokument 'd'.
- $nDok(g)$: Anzahl der Dokumente, die eine Grundform 'g' enthalten.
- $nColl(g)$: Anzahl der Vorkommen, der Grundform 'g' in der Dokumentenkollektion.
- $nGruLen(g)$: Länge der Grundform 'g'.

Mit den oben definierten Parametern, die von SELIX für bestimmte Grundformen (Indexkandidaten) berechnet werden, lassen sich auch Gewichtsfunktionen aus der Literatur benutzen. Für das System SELIX wurden hierzu zwei Funktionen ausgewählt:

Salton (SA 89)

$$\text{HfklmD}(g,d) * \log (n\text{Dok}(g) / n\text{AnzDok})$$

Robertson (HR 98)

$$((K + 1) * \text{HfklmD}(g,d) / (K + \text{HfklmD}(g,d))) * \log ((n\text{AnzDok} - n\text{Dok}(g) + 0.5) / (n\text{Dok}(g) + 0.5))$$

Unabhängig von obigen beiden Gewichtungsfunktionen und weiteren Funktionen aus der Literatur wurden zwei eigene Gewichtungen entwickelt. Dabei haben (wie schon erwähnt) drei Gesichtspunkte eine wichtige Rolle gespielt. Für jeden Gesichtspunkt wurde ein Teilgewicht entwickelt.

Mit Hilfe obiger Zahlen werden zu jedem Indextermkandidat die drei Teilgewichte

nG1 (kollektionsbezogen (für alle Dokumente (einer Kollektion) gleich),

nG2 (dokumentbezogen) und

nG3 (unabhängig von einem Dokument und einer Kollektion)

ermittelt und additiv miteinander zu dem Gesamtgewicht nG verknüpft. Dabei können die Gewichte noch mit einem Faktor versehen werden (F1 – F3).

$$nG = F1 * nG1 + F2 * nG2 + F3 * nG3$$

Da aus der Literatur kein Ansatz bekannt war, der diese drei Gesichtspunkte berücksichtigt, wurde diese Vorgehensweise gewählt. Es ist denkbar, dass für die Teilgewichte nach den Retrievaltests auch Teile der Funktionen bei Salton, Robertson oder anderen aus der Literatur bekannten Funktionen genommen werden, bzw. dass die 'eigenen' verbessert werden.

Jedes Teilgewicht (nG1 – nG3) soll in dem Intervall (0,1) liegen, damit die Benutzung verschiedener Faktoren (F1 – F3) übersichtlicher wird.

4.1 Das Kollektionsgewicht nG1

In dem Teilgewicht nG1 soll sich widerspiegeln, ob eine Grundform für eine Dokumentenkollektion als Indexterm geeignet ist. Hierbei wird sich an einer Zufallsverteilung orientiert (siehe hierzu neben (SA 89), (HR 98) auch (HA 75)). Aus der Überlegung 'wenn sich viele Vorkommen einer Grundform auf relativ wenige Dokumente konzentrieren, kann keine zufällige Verteilung vorliegen' wurde für **KasEinf** folgende Funktion gewählt.

KasEinf (KASCADE einfach):

$$nG1 = 1 - n\text{Dok}(g) / n\text{Coll}(g)$$

Bei **KasKomp** wurde für $nDok(g)$ als Vergleichsgröße der Erwartungswert $E(nDok(g))$ für die Anzahl der Dokumente, in denen die Grundform 'g' vorkommt, genommen.

KasKomp (KASCADE komplex):

$$nG1 = 1 - nDok(g) / E(nDok(g)) \quad (\text{für: } nDok(g) < E(nDok(g)); 0 \text{ sonst})$$

mit

$$E(nDok(g)) = nAnzDok * (1 - \exp(-\lambda))$$

wenn eine Poisson Zufallsverteilung angenommen wird:

$$P(i) = \exp(-\lambda) * (\lambda^i / i!) \quad \lambda = nColl(g) / nAnzDok$$

4.2 Das Dokumentgewicht nG2

In dem Teilgewicht $nG2$ soll sich widerspiegeln, ob eine Grundform für eine Dokument als Indexterm wichtig ist.

Aus der Überlegung 'wenn die Häufigkeit einer Grundform 'g' innerhalb eines Dokumentes 'd' im Vergleich zur durchschnittlichen Häufigkeit groß ist, muss diese Grundform für dieses Dokument wichtig sein' wurde für **KasEinf** folgende Funktion gewählt.

KasEinf:

$$nG2 = 1 - ((nDokLen(d) / nCollen) / (HfklmD(g,d) / nColl(g)))$$

(falls ≥ 0 ; sonst 0)

Bei **KasKomp** wurde als Gewicht der Erwartungswert, dass der Indexterm 'g' in dem Dokument 'd' 1 bis $HfklmD(g,d)$ mal vorkommt, in Relation zum Erwartungswert gewählt.

KasKomp:

$$nG2 = (p(1) * 1 + \dots + p(HfklmD(g,d)) * HfklmD(g,d)) / \lambda$$

mit

$$P(i) = \exp(-\lambda) * (\lambda^i / i!)$$

$$\lambda = nColl * (nDokLen / nCollen)$$

4.3 Das Längengewicht nG3:

Betrachtet man Begriffe wie

EU-Denkmalförderungsrichtlinie,
 Elektrizitätsbinnenmarktrichtlinie,
 Urheberrechtsschiedsstellenverordnung oder
 Schiedsgerichtsvereinbarungen in Rückversicherungsverträgen,

dann sind diese Terme sicherlich wichtig für ein Dokument, in denen sie vorkommen, ebenso stehen sie kaum zufällig in einem Dokument, etwa im Vergleich zu 'Richtlinie'. Wenn jedoch keine sehr große Dokumentenkollektion vorliegt, und solche Terme selten vorkommen, haben diese Terme bei einem statistischen Ansatz kaum eine Chance. Deshalb wurde bei dem dritten Gewicht nG3 die Länge miteinbezogen.

$$nG3 = \log (nGruLen(g)) / 4$$

5 Das Programmsystem von SELIX

Durch Auswahl bestimmter Wortklassen, 'Relationen' und 'Semantiken' (siehe (LZ 97)) können bestimmte Grundformen von der Häufigkeitszählung ausgeschlossen werden, bzw. können verschiedene Grundformen auf eine Vorzugsgrundform bei der Zählung zusammengeführt werden.

Bei Komposita und Mehrwortgruppen die Möglichkeit der Mehrfachzählungen (zusätzlich zum Teilgewicht nG3).

Zur Zeit werden die Gewichte 'nG' nur benutzt, um Indexterms für ein Dokument zu selektieren. Die Länge eines Dokumentes soll dabei auch für die Anzahl der Deskriptoren berücksichtigt werden. Hierzu gibt es folgende Parameter:

nGewMin: Minimales Gewicht.

nAnzMin: Minimale Anzahl von Deskriptoren.

nAnzMax: Maximale Anzahl von Deskriptoren.

nAnzMit: Mittlere maximale Anzahl von Deskriptoren. Diese Anzahl gilt für ein durchschnittlich langes Dokument (Länge: nColLen(d) / nAnzDok). Für jedes Dokument wird diese Anzahl (Mittlere maximale Anzahl von Deskriptoren) in Abhängigkeit von seiner Länge neu berechnet. Diese Zahl muss aber zwischen den Zahlen nAnzMin und nAnzMax liegen, die unabhängig von der Dokumentlänge sind.

Wenn eine Grundform als Registerstichwort gekennzeichnet ist, wird diese Grundform ausgegeben, unabhängig von seinem Gewicht. Zusätzlich wird bei dem Gewicht das Zeichen '*' vorangestellt.

Neben einem 'formalen Trivialwort' (z.B. 'der') gibt es noch ein 'Trivialwort (System)' (z.B. 'Kapitel') und ein 'Trivialwort (Nutzer)' (z.B. 'Recht' im Fachgebiet Jura). Ein formales Trivialwort wird als Indexterm nicht zugelassen. Die andern beiden werden bei der Gewichtsrechnung behandelt wie andere Indexkandidaten; bei der Ausgabe wird, falls sie die anderen Bedingungen erfüllen, dem Gewicht die Zeichenkette '*' vorangesetzt.

6 Ausblick

In dem System SELIX ist schon die Möglichkeit zur automatischen Gewinnung von Mehrwortgruppen (mit Hilfe eines MWG-Parsers) vorgesehen.

Beispiele aus ersten Tests:

Diskrepanz zwischen der ermittelten Blutalkoholkonzentration und Zeugenaussage

Ermittlung der Blutalkoholkonzentration aus der Blutprobe / ohne Blutuntersuchung

fahrlässige Herbeiführung der Beeinträchtigung der Schuldfähigkeit

Eine solche Mehrwortgruppe ist sicher nicht geeignet als Indextermkandidat im herkömmlichen Sinne. Diese Mehrwortgruppen sollen jedoch (in Zukunft) benutzt werden bei der Dialogführung mit dem Benutzer, wenn dieser mit bestimmten Teilkomponenten solcher Mehrwortgruppen in den Retrievaldialog einsteigt.

7 Literatur

[HA 75] Harter, S.D. (1975): A probabilistic approach to automatic keyword indexing (On the distribution of specialty words in a technical literature); Journal of ASIS, (1975), p. 197-206

[HR 98] Huang, Xiangji; Robertson, S.E. (1998): Okapi Chinese text retrieval experiments at TREC-6

[LZ 97] Lepsky, Klaus; Zimmermann, Harald H. (1997): Katalogenerweiterung durch Scanning und Automatische Dokumenterschließung; Das DFG-Projekt KASCADE
http://www.uni-duesseldorf.de/WWW/ulb/kas_home.htm

[SA 89] Salton, G. (1989): Automatic Text Processing; Addison Wesley

Team Finder

Ein intelligenter Assistent zur zielgerichteten Ressourcenselektion für workflow-integrierte multimediale Audio- / Video- Desktopkonferenzen

Ruppert von Teutul, Georg Schneider, Jean Schweitzer

Ruppert von Teutul, DFKI GmbH, Saarbrücken
Georg Schneider, DFKI GmbH, Saarbrücken
Jean Schweitzer, Siemens AG, Saarbrücken
Siemens Telekooperationszentrum am DFKI
Stuhlsatzenhausweg 3
D-66123 Saarbrücken, Germany

Inhalt

- 1 Inhalt
- 2 Zusammenfassung
- 3 Einleitung
- 4 Integration
- 5 Aufgaben des Team Finder
- 5.1 Unterstützung bei der Selektion der Teilnehmer
- 5.2 Bestimmung eines gemeinsamen Termins
- 5.3 Transparente Integration in das Workflow-Management-System
- 6 Teambewertungen
- 7 Realisierung der Selektion
- 8 Ausblick
- 9 Literatur

2 Zusammenfassung

Team Finder unterstützt die Bildung von Teams u.a. durch Angabe von Kompetenz- und Rollenspezifikationen und ihre Einbindung in Workflow-Management-Systeme als Einheit. *Team Finder* findet Konferenztermine selbständig und vermag durch Berücksichtigung der Historie und Leistungsfähigkeit der gebildeten Teams, herausragende Teams bei der Selektion zu bevorzugen. Die

dazu notwendigen Bemessungen erfolgen mittels sozial- und arbeitspsychologischer Fragebögen am Ende einer jeden Konferenz.

3 Einleitung

Workflow-Management-Systeme gewinnen im Zuge der Optimierung und Verbesserung von Arbeitsabläufen immer stärker an Bedeutung. Sie helfen nicht nur bei der Strukturierung der Arbeitsprozesse in Organisationen, sie gestalten Arbeitsabläufe auch klarer und reproduzierbar. Durch den Einsatz von Workflow-Management-Systemen kann die korrekte Abarbeitung dieser Vorgänge garantiert und beschleunigt werden. Zu jedem Zeitpunkt sind Status und Bearbeiter genau feststellbar. Diese workflowgestützte asynchrone Kooperation um computergestützte synchrone Kooperation zu erweitern, ist Bestandteil des Projektes POLIVEST [Dietel, Schneider, Schweitzer 97]. Die dort betrachteten Szenarien kommen aus dem Umfeld des Bundesrates, bzw. der Bundes- und Landesämter. Sie bieten für den Einsatz von Workflow-Management-Systemen hinreichend komplexe Abläufe, wie beispielsweise die Bearbeitung von Gesetzesvorlagen im Bundesrat oder von Bauanträgen im Rhein-Sieg-Kreis. Weiterhin sind diese Prozesse durch hohe Anforderungen hinsichtlich zeitgleicher Kooperation geprägt, wie die gemeinsame Besprechung der Gesetzesvorlagen zwischen Landes- und Bundesverwaltungen oder der Besprechung eines Bauvorhabens zwischen Baubehörde, Bauherr und Architekturbüro. Durch den Einsatz einer ganzheitlichen, rechnergestützten Kooperationsumgebung werden Reaktions- und Bearbeitungszeiten verkürzt, und die Zahl der unerwünschten Medienbrüche von Papier zu elektronischen Dokumenten reduziert [Schneider, Baumann, Schweitzer 97]. In POLIVEST wird die Kooperation via Multimedia Audio-/Video-Desktopkonferenzen (MMC) durch das System GroupWin bereits realisiert [GroupWin 96].

Durch die Integration dieser MMC Systeme in Workflow-Management-Systeme bilden sich kooperierende und gemeinsame Aufgaben bearbeitende Gruppen, die allerdings bisher von den Workflow-Management-Systemen nicht als solche erfaßt werden können. Wenn im oben beschriebenen Umfeld gemeinsam eine Entscheidung getroffen werden soll, müssen Workflow-Management-Systeme eine Möglichkeit eröffnen, mehrere Teilnehmer für die Arbeit als Team zusammenzufassen. Eben solche Anforderungen entstehen auch beim Auftreten von ad-hoc Konferenzen, zum Beispiel zur Lösung unvorhergesehener Probleme. In heutigen Systemen ist es zwar möglich, Projektgruppen zu definieren, sie ermöglichen aber nicht, mehrere Mitarbeiter gleichzeitig an einer Task innerhalb eines Workflows arbeiten zu lassen. Das Konzept der zeitgleichen Teamarbeit ist in Workflow-Management-Systemen nicht verfügbar, obwohl dieser Form von Teamarbeit in modernen Unternehmen wachsende Bedeutung zukommt.

Andererseits bieten Workflow-Management-Systeme entsprechend der Workflow Management Coalition Komponenten, die den Zugriff auf die Aufbauorganisation von Unternehmen zulassen [WfMC 98], was umgekehrt bei MMC Systemen nicht möglich ist. Damit können einzelne Aufgaben innerhalb eines Workflows nicht nur direkt an Mitarbeiter, sondern auch an bestimmte Kompetenzen, Stellen oder Rollen zugewiesen werden. Ein weiterführender Ansatz ist, diese Konzepte bei der Bildung von Teams ebenfalls zuzulassen.

Ansätze zur Integration von Workflow-Management- und MMC Systemen, wie beispielsweise Realtime Notes [Lotus 98], ermöglichen nur den gemeinsamen Zugriff auf ein unternehmensweites Adreßbuch eines Unternehmens. Systeme aus dem universitären Umfeld wie Planko [Schaller, Schwab 97], WAM [Messer, Faustmann, 95] oder Contact [Galler, Hagemeyer, Scheer 95] beziehen sich ausschließlich auf die Koordination von Arbeitsabläufen.

Unser Ziel bei *Team Finder* ist die intelligente Selektion von Teams mit besonderem Augenmerk auf ihre Qualität im Sinne effizienter Teamarbeit [Schneider, Schweitzer 97]. So kann in *Team Finder* die Leistungsfähigkeit eines Teams bewertet oder die Kompetenzen einzelner Mitarbeiter in bezug auf einen Arbeitsablauf ermittelt werden. Neben den formalen Anforderungsprofilen wie Kompetenzen, Stellen oder Rollen, können zusätzlich auch oben genannte Einflußfaktoren in den Auswahlprozeß einbezogen werden. Dadurch wird erreicht, daß besonders erfolgreiche und leistungsfähige Teams wiederholt für gleiche oder ähnliche Aufgaben herangezogen werden.

Diesen Anforderungen entsprechend, liefert das von uns entwickelte System *Team Finder* einen Ansatz zur Integration von Teams in die Aufbauorganisation und die Integration von Teamarbeit via Multimedia Audio-/Video-Desktopkonferenzen in die workflowunterstützte Vorgangsbearbeitung [Dietel, Schneider, Schweitzer 98]. Damit wird mit dem Workflow Paradigma „one person – one task – one application“ gebrochen [Stark, Lachal 95]. Weiterhin werden leistungsfähige Teams eines Unternehmens ermittelt und für spätere Referenz abgelegt.

4 Integration

Die Bildung von Teams spielt in Unternehmen eine immer wichtigere Rolle. Sie sorgt für eine bessere Nutzung von Ressourcen und schafft hochwertigere und kreativere Problemlösungen [Rosenstiel 92]. *Team Finder* unterstützt Workflow-Designer und Sachbearbeiter bei der Selektion der Konferenzteilnehmer für die MMC und erweitert damit das eingeschränkte Selektions- und Leistungsspektrum herkömmlicher Workflow-Management-Systeme um den Faktor Teamarbeit.

Eine Task eines Workflows ist beispielsweise die Durchführung einer MMC. Diese MMC hat mehrere Teilnehmer, die interaktiv miteinander agieren müssen, z.B. indem sie gemeinsam ein Dokument bearbeiten. Diese Teilnehmer können Personen sein, die in dem Workflow mitarbeiten, sie können auch Personen sein, die dort nicht mitarbeiten und solche, die gar nicht an das Workflow-Management-System angeschlossen sind, wie z.B. externe Berater. Sie müssen jedoch alle in der Organisationsstruktur abgelegt sein und stehen dadurch zur Auswahl durch den *Team Finder* bereit. Nach Beendigung der MMC wird der Workflow weiter ausgeführt.

Eine konferenzvorbereitende Aktivität des Workflows bildet die Bestimmung der Teilnehmer durch einen Verantwortlichen. Dabei können verschiedene Kriterien Anwendung finden; das Ergebnis dieser Selektion ist eine Menge von Gruppen von potentiellen Teilnehmern, die zu bestimmten Terminen zur Verfügung stehen. So können durch diese Kriterien ebenso Teams gebildet werden, die zum schnellstmöglichen Zeitpunkt in einer Konferenz teilnehmen können, wie auch

Teams, die bereits vorher gut miteinander gearbeitet haben oder besonders genau passende Kompetenzen besitzen.

Team Finder legt dann den Termin fest und bringt die Konferenz zum festgelegten Zeitpunkt automatisch mit den geladenen, selektierten Teilnehmern zur Ausführung.

5 Aufgaben des Team Finder

Team Finder erfüllt zur Konferenzvorbereitung drei Hauptaufgaben

- Unterstützung bei der Selektion der Teilnehmer
- Bestimmung eines gemeinsamen Termins
- Transparente Integration in das Workflow-Management-System

5.1 Unterstützung bei der Selektion der Teilnehmer

Bei der Selektion der Teilnehmer werden die Daten der Aufbauorganisation zugrunde gelegt. Bei *Team Finder* handelt es sich um eine unabhängige, generische Lösung, die vom zugrundeliegenden Workflow-Management-System abstrahiert. Deshalb läßt sich *Team Finder* in die strukturellen Gegebenheiten unterschiedlicher Workflow-Management-Systeme integrieren. Das System *Team Finder* bietet dem Verantwortlichen die in der Aufbauorganisation abgelegten Konzepte (Objektklassen) wie Rolle, Stelle, Kompetenz, Mitarbeiter usw. zur Selektion der potentiellen Teammitglieder an.

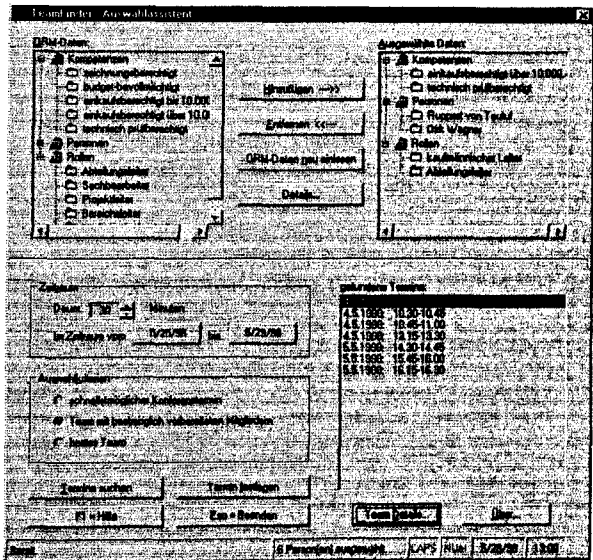


Abb. 1: Der Auswahl-Assistent des Team Finder

Die aus dieser Voreinschränkung resultierende Menge von Mitarbeitern wird durch qualifizierende Nebenbedingungen weiter eingeschränkt.

5.2 Bestimmung eines gemeinsamen Termins

Ein wesentlicher Teil bei der Bestimmung eines gemeinsamen Termins ist die Einschränkung durch die angegebenen Kriterien. *Team Finder* unterstützt fol-

gende Kriterien, die gemeinsam Anwendung finden können und dabei in ihrer Gewichtung variieren können.

R1. *schnellstmöglicher Konferenztermin*

Die Konferenzteilnehmer sollen so aus der Gesamtheit der möglichen Teilnehmer selektiert werden, daß sie den spezifizierten Rollen, Stellen oder Kompetenzen entspringen und die Konferenz zum schnellstmöglichen Zeitpunkt durchgeführt werden kann.

Organisations- und sozialpsychologische Aspekte der Konferenzplanung bleiben bei diesem Kriterium allerdings unberücksichtigt.

Um diese Perspektive erweitert, ist das folgende Kriterium von Interesse:

R2. *Team mit bestmöglich vorbereiteten Mitgliedern*

Dieses Team zeichnet sich durch die Kompetenzen der Teilnehmer aus. Es wird den Teilnehmern Vorrang gegeben, die bereits im betreffenden Workflow mitgearbeitet haben und somit eine besonders hohe Kompetenz bzgl. dieses Workflows besitzen.

Der Vorteil dieses Kriteriums liegt darin, daß die Teilnehmer mit einem Minimum an zusätzlicher Information über den in der Konferenz zu diskutierenden Sachverhalt ausgestattet werden müssen.

Die Teilnehmer können nicht nur nach ihren formalen und fachlichen Fähigkeiten ausgewählt werden; es wird auch ihre Kooperationsfähigkeit mit anderen Teilnehmern berücksichtigt. Dadurch können langwierige Verhandlungen vermieden werden, die den Ablauf von Konferenzen beeinflussen und die Bearbeitung des Workflows verzögern.

Um diese Störungen zu verhindern, werden die Bewertungsergebnisse aus vorherigen Konferenzen zu einem Team-Ranking zusammengefaßt. Auf dieses Ranking kann mittels des folgenden Kriteriums zugegriffen werden:

R3. *bestes Team*

Dieses Team qualifiziert sich durch gute Zusammenarbeit der Teammitglieder und schnellen bzw. effizienten Konferenzablauf. Es kann davon ausgegangen werden, daß die Teilnehmer bei gleicher Problemstellung erneut ein gutes und schnelles Konferenzergebnis erzielen.

Die Bewertung des Konferenzverlaufes wird durch Befragung der einzelnen Teilnehmer am Ende der Konferenz bestimmt und in einer tabellenähnlichen Struktur hinterlegt. Auf die Realisierung der Bewertung wird in Abschnitt 6 genauer eingegangen.

Jedes der drei aufgeführten Kriterien kann mit einer bestimmbaren Gewichtung in den Selektionsprozess aufgenommen werden. Bei Anwendung des Kriteriums R1 werden R2 und R3 nicht bewertet. Bei R2 und R3 hingegen, wird R1 implizit immer angewendet, denn das erste Kriterium hat bei Anwendung der beiden anderen immer niedrigste Priorität. Die Realisierung der Selektion wird in Abschnitt 7 genauer beschrieben.

5.3 Transparente Integration in das Workflow-Management-System

Der Bearbeiter gibt für die Konferenz einen Zeitraum und eine voraussichtliche Dauer an. *Team Finder* ermittelt die den Konzepten Rolle, Kompetenz usw. entsprechenden Teilnehmer und stellt sie zu Teams mit den gewünschten Eigenschaften (Kriterien) zusammen. Als Ergebnis schlägt *Team Finder* dem Benutzer geordnete, passende Kombinationen von Teams und Terminen vor, zu denen alle Teammitglieder Zeit haben. Der Bearbeiter wählt einen Termin mit dem dazu gehörenden Team, das seinen Vorstellungen entspricht. Dieser Termin wird im Workflow-Management-System und in den Terminkalendern der Teilnehmer eingetragen. Automatisch wird im Verlauf des Workflows zum gegebenen Zeitpunkt die MMC Konferenz gestartet.

6 Teambewertungen

Bei der Teambewertung wird auf ein Team-Ranking zurückgegriffen. Mehrere Verfahren für die Bildung dieses Rankings sind denkbar. So kann z.B. anhand der abgearbeiteten Konferenzpunkte unter Einbeziehung der Konferenzdauer auf die Effizienz der Teamarbeit geschlossen werden. Durch die Befragung der Teilnehmer selbst, mittels eines WEB-basierten, elektronischen Fragebogens, der einer Einzelbefragung recht nahe kommt, kann die Effizienz der Teamarbeit ebenfalls gemessen werden. Diese Verfahren führen zu einem Team-Ranking, auf das später bei der Anwendung des Kriteriums „bestes Team“ zurückgegriffen wird.

Team Finder stellt zur Ermittlung des Rankings mehrere aufgabenbezogenen Fragebögen zur Auswahl [Francis, Young 96]. Der vom Bearbeiter bei der Konferenzvorbereitung ausgewählte Fragebogen wird nach Ablauf der MMC an alle Teilnehmer verschickt und nach Beantwortung durch das System ausgewertet. Die gewonnenen Resultate werden im System tabellenartig abgelegt und zur Auswahl folgender Teams in folgenden Workflows herangezogen. Diese Auswertung wird den Teammitgliedern ebenfalls zugestellt, evtl. mit daraus resultierenden Hinweisen zur Verbesserung der Teamarbeit. Im Laufe der Zeit wird so eine Verbesserung der Teamarbeit innerhalb einzelner Teams erreicht. Dies ist zusätzlich ein Beitrag zur lernenden Organisation [Senge 92].

Die verwendeten Fragebögen erfassen Daten, die Rückschlüsse auf die Kombination der Teilnehmer, also das Team, nicht jedoch auf die einzelnen Teilnehmer selbst, zulassen. Die Fragen gliedern sich in mehrere Bereiche.

Durch dieses flexible und generische System bietet *Team Finder* dem Gestalter der WEB Abfragen genügend Spielraum, um den unterschiedlichen Aufgabenstellungen entsprechend, Fragebögen zu adaptieren, weitere Fragebögen zu formulieren oder auch andere Bewertungsmechanismen zu integrieren, wie z.B. die Bewertung durch ein Komitee.

7 Realisierung der Selektion

Um die gewonnenen Daten effizient und intelligent auswerten zu können, werden sie für die Bildung neuerlicher Teams dem Constraintsolver Oz zur Verfü-

gung gestellt [Smolka, Treinen 95]. Er erhält als Eingabe die Menge der vorselektierten Teilnehmer sowie die geforderten Profile. Die Angabe des Selektionskriteriums führt zusammen mit den bereits vorher gewonnenen Rankings zur Bildung von Ausgaben der Form: (Team / Konferenztermin). Würde z.B. das beste Team erst in mehreren Wochen einen Termin frei haben und das zweitbeste Team einen Termin in wenigen Tagen, garantiert *Team Finder*, daß das zweitbeste Team in der Liste der Empfehlungen ganz oben steht. Durch Einsatz reifizierter Constraints wird ermöglicht, die Kriterien R1 bis R3 selektiv in die Auswertung einzubeziehen. Das bei Anwendung von R1 bis R3 entstehende Problem ist ähnlich dem Scheduling und Resource Allocation Problem (siehe beispielsweise [Henz, Würtz 95][Würtz 97]). Dabei werden jeweils für die Kriterien R1 bis R3 die Constraints zusammengefaßt und mit unterschiedlichen Gewichtungen auf die Eingabemenge der Personen mit ihren Qualifikationen angewendet. Oz ermittelt für diese Einschränkungen dann die Resultatsmenge.

8 Ausblick

Unternehmen werden in Zukunft zur Kommunikation verstärkt zu Mitteln aus dem Bereich MMC greifen. *Team Finder* kann in diesem Umfeld zur Integration von Telearbeit (siehe z.B. [Hoechst, Schneider, Merschjan, Schweitzer 98]) in bestehende Workflow-Management-Systeme eingesetzt werden und helfen, zur Verwirklichung virtueller Teams im Sinne ganzheitlicher Prozeßbearbeitung beizutragen.

Team Finder wird im Projekt POLIVEST Verwendung finden. Die Erprobung in diesem Umfeld liefert erste Ergebnisse und Erfahrungen, die Rückschlüsse auf notwendige Erweiterungen des Systems zulassen.

Knowledge-Management, welches bisher nur auf die Kompetenzen von Einzelpersonen abzielt, kann durch *Team Finder* nun um Gruppen und Teams erweitert werden. Damit eröffnen sich neue Möglichkeiten, die Leistungsfähigkeit von Unternehmen zu überblicken und zu steigern.

Team Finder kann außerdem als Testumgebung zur Entwicklung von Meßverfahren für die Team-Effizienz eingesetzt werden. Durch seine Modularität kann es leicht unterschiedlichen Gegebenheiten angepaßt werden. Weiterführende Arbeiten können z.B. die Funktion des Moderators oder die Performance einzelner Mitglieder berücksichtigen und somit zu noch effizienteren Teambildungen führen.

9 Literatur

[Dietel, Schneider, Schweitzer 97]

C. Dietel, G. Schneider, J. Schweitzer: *POLIVEST Integrierte Televerwaltung*, in: Tagungsband Workshop: Rechnergestützte Kooperation in Verwaltungen und großen Unternehmen, im Rahmen der 27. Jahrestagung der Gesellschaft für Informatik (Informatik '97), Aachen, 24.-27. September 1997, S.188-205

[Dietel, Schneider, Schweitzer 98]

C. Dietel, G. Schneider, J. Schweitzer: *A Team Oriented Approach Towards the Integration of Workflow Management and Multimedia AV Desktop Conference Systems in the Context of Public Administrations*, in: Proceedings of the 36th International Conference on Telecommunication Systems, Modelling and Analysis, Nashville, Tennessee, USA, March 5-8, 1998

[Francis, Young 96]

D. Francis, D. Young: *Mehr Erfolg im Team, Ein Trainingsprogramm mit 46 Übungen zur Verbesserung der Leistungsfähigkeit in Arbeitsgruppen*, Windmühle GmbH, Hamburg, 1996

[Galler, Hagemeyer, Scheer 95]

J. Galler, J. Hagemeyer, A-W. Scheer: *ContAct - Coordination of Cooperative Information Modeling Activities*, in: Conference Supplement to the European Computer Supported Cooperative Work, Stockholm, Sweden, 1995, p.25-26

[GroupWin 96]

GroupWin: *Benutzerhandbuch*, Siemens Telekooperationszentrum am DFKI, 66123 Saarbrücken, 1996

[Henz, Würtz 95]

M. Henz, J. Würtz: *Using Oz for College Timetabling*, in: The Practice and Theory of Automated Time Tabling: The Selected Proceedings of the 1st International Conference on the Practice and Theory of Automated Time Tabling, Edinburgh, 1995

[Hoechst, Schneider, Merschjan, Schweitzer 98]

P. Hoechst, G. Schneider, K. Merschjan, J. Schweitzer: *Telearbeit auf der Grundlage einer neuen Kommunikationstechnologie*, in: Tagungsband, Sixth International Symposium for Information Science (ISI '98), 1998

[Lotus 98]

Lotus Development Corp: <http://www.lotus.com/>, 1998

[Schaller, Schwab 97]

T. Schaller, K. Schwab: *Integration von asynchronen CSCW-Anwendungen*, in: S. Jablonski, M. Böhm, W. Schulze, (Hrsg.), *Workflow-Management - Entwicklung von Anwendungen und Systemen - Facetten einer neuen Technologie*, dpunkt.verlag Heidelberg, 1997, S.374-399

[Schneider, Baumann, Schweitzer 97]

G. Schneider, S. Baumann, J. Schweitzer: *Integrating and coordinating major software systems with workflow management systems*, in: D. T. Wright, M. M. Rudolph, V. Hanna, D. Gillingwater, N. D. Burns (Eds.), *Proc. Managing*

Enterprises - Stakeholders, Engineering, Logistics and Achievement (ME-SELA'97), Loughborough University, UK, Mechanical Engineering Publications Limited, London and Bury St Edmunds UK, July 22-24, 1997

[Schneider, Schweitzer 97]

G. Schneider, J. Schweitzer: *Closing the gap between synchronous and asynchronous cooperative work*, in Proc. Fifteenth International Joint Conference on Artificial Intelligence (IJCAI 97), Workshop „Business Applications of AI“, NAGOYA, Aichi, Japan, August 23-29, 1997

[Smolka, Treinen 95]

G. Smolka, R. Treinen: *DFKI Oz Documentation Series*, Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) GmbH, Saarbrücken, 1995

[Senge 92]

P. Senge, Building Learning Organizations: *The Journal for Quality & Participation*, Association for Quality and Participation, Cincinnati, Ohio, <http://www.aqp.org/journal.html>, März 1992

[Stark, Lachal 95]

H. Stark, L. Lachal: *Ovum Evaluates Workflow*, Ovum Ltd., London, 1995

[Messer, Faustmann 95]

B. Messer, G. Faustmann: *Effiziente Videokonferenzen durch Workflow-Management-Systeme*, in: K. Hammer, D. Schmolke, F. Stuchlik (Hrsg.), Tagungsband Synergie durch Netze, Otto-von-Guericke-Universität Magdeburg, 5.-6. Oktober, 1995

[WfMC 98]

Workflow Management Coalition: *Referenzmodell*, http://www.aiim.org/wfmc/IMG/ref_model.html, 1998

[Würtz 97]

J. Würtz: *Constraint-Based Scheduling in Oz*, Operations Research Proceedings 1996, Springer Verlag, 1997

Text-Fakten-Integration in Informationssystemen

Jürgen Krause, Thomas Mandl, André Schaefer,
Maximilian Stempfhuber

IZ Sozialwissenschaften, Lennéstr. 30, D-53113 Bonn, st@bonn.iz-soz.de
Universität Koblenz-Landau, Institut für Informatik

Inhalt

- 1 Text-Fakten-Integration
- 2 Das Verbandsinformationssystem ELVIRA
- 3 Konzept für die Text-Fakten-Integration in ELVIRA
 - 3.1 Text- und Faktensuche in Reinform
 - 3.2 Transformationen zwischen den Suchtypen
 - 3.3 Modalitätsunabhängige Suche
 - 3.4 Zusammenfassung und weiteres Vorgehen
- 4 Textrecherche mit integrierter Faktensuche
- Literatur

Zusammenfassung

Der Beitrag stellt den Prototypen eines integrierten Informationssystems vor, das Daten verschiedener Typen (Modalitäten) enthält. Das zugrundeliegende Modell entstand aus empirischen Untersuchungen für Texte und Fakten, eignet sich aufgrund der gewählten Abstraktionsstufe jedoch auch für andere Datentypen wie z.B. Grafiken oder Internet-Seiten. Als Anwendungsfall dient das Elektronische Verbandsinformations-, Recherche- und Analysesystem ELVIRA.

1 Text-Fakten-Integration

Heterogenitätsbehandlung im Information Retrieval ist eine der großen Herausforderungen für zukünftige Informationssysteme. Ein Spezialfall hiervon ist die integrierte Suche nach Texten und Fakten, die bei zahlreichen Dokumenttypen sinnvoll sein kann (z.B. bibliographische Daten, Faktendatenbanken, Datenbanken für Forschungsprojekte, Dienstleistungsangebote, Rechner, Förderprogramme).

Bei der Text-Faktenintegration können drei Typen von Suchen vorkommen:

a) Kombinierte Suchen

Bei kombinierten Suchen werden Bedingungen für Texte und Fakten und ihre Zusammenhänge in einer Query formuliert, wodurch die Beziehungen zwischen den Dokumenten besser ausgenutzt werden. Ein Beispiel ist die Faktensuche bei der man sich primär auf die Dokumentattribute (i.d.R. bibliographische Angaben) bezieht. Dies ist zum einen zur Realisierung bestimmter Suchstrategien im Rahmen einer Dokumentrecherche notwendig, etwa wenn man nach Autoren, Zeitschriften oder Konferenzen zu einem bestimmten Thema sucht, um mit diesem Ergebnis die Suche voranzutreiben. Zum anderen sind die Fakten selbst häufig das Ziel einer Suche, wenn man z.B. nach Autoren als Experten auf einem bestimmten Gebiet sucht oder nach Institutionen, die als Kooperationspartner in Frage kommen.

b) Umstiege durch Transformationen

In vielen Situationen wollen Nutzer z.B. nach einer Textanfrage mehr zu einem Thema wissen und möchten nun auch Fakten und z.B. Internet-Seiten dazu sehen. Da die Inhaltserschließung bei diesen verschiedenen Modalitäten unterschiedlich ist und oft anderes Vokabular benutzt wird, müssen bereits formalisierte Informationsbedürfnisse in Anfragen der anderen Modalitäten transformiert werden. Transformationen sollten in allen Richtungen (von Fakten zu Texten, von Texten zu Fakten, usw.) und sowohl aus initialer Suchanfrage als auch aus dem Ergebnis heraus möglich sein.

c) Darstellungsform-unabhängige Anfragen

Informationsbedürfnisse können auch von vorne herein unabhängig von der Darstellungsform des gespeicherten Wissens sein, also sowohl durch Textdokumente als auch durch Faktendaten beantwortet werden. Z.B. kann eine Information über eine Chemikalie in einer Tabelle oder einem beschreibenden Text gefunden werden. Für solche Informationsbedürfnisse muß ein System die Formulierung einer generellen Anfrage zulassen, diese in alle modalitätsabhängigen Anfragen transformieren und die Ergebnisse kombinieren.

Obwohl das Bewußtsein, daß bei einer realen Entscheidungs- oder Problemlösungssituation verschiedene Dokumenttypen, vor allem auch unstrukturierte Dokumentbestände (Texte) neben strukturierten wichtig sind, weit verbreitet ist; beschränkt sich ihre Integration in der Regel auf die Fragen der technischen Verwaltung in verteilten Datenbanken und die technisch effiziente Verarbeitung von Anfragen (z.B. [Gu et al. 1993]). Darüberhinaus wird die Integration von Texten und Fakten oft als erster Schritt einer multimedialen Verknüpfung verstanden (cf. [Womser-Hacker 1997]).

Praxisbeispiele einer Integration von Text- und Fakten-Material sind die Anwendungen einer schweizerischen Großbank (cf. [Amstutz et. al. 1991]) und ein Kriminalinformationssystem der niederländischen Polizei (cf. [Hoogeveen et al. 1992]). Ein Konzept für kombinierte Suchen, bei denen sowohl Bedingungen für Fakten als auch für Texte formuliert werden, bietet [Fuhr 1992]. Auf diesen Typ beschränken sich auch viele Anwendungen, wie etwa die bei der GMD-IPSI entwickelte experimentelle Forschungsdatenbank HYDRA (cf. [Gu et al. 1993]), die Fakten- und Textwissen über EU-Projekte beinhaltet. Die beschreibenden Textteile sind verbunden mit strukturierten Fakteninformationen wie dem Projektleiter, Förderstelle oder Projektlaufzeit.

Erste konzeptuelle Grundlagen für eine Integration von Zeitreihen und Texten enthält für die Sozialwissenschaften Marx et al. 1995 und für die wirtschaftlichen Zeitreihen des Verbandsinformationssystems ELVIRA Krause et al. 1997. Zu diesen Vorüberlegungen gehört auch das MIMOR-Modell von Womser-Hacker 1997, das die Text-Fakten-Integration in einem System der Mehrfachindexierung zur dynamischen Methoden-Objekt-Relationierung im Information Retrieval berücksichtigen möchte. Methoden für Transformationen diskutiert [Mandl 1998].

Das Konzept für die Text-Fakten-Integration im Projekt ELVIRA II (vom Bundesministerium für Wirtschaft gefördert unter Aktenzeichen IV C2-003060/22), das im folgenden Überblick dargestellt wird, basiert auf empirischen Untersuchungen mit Nutzern von ELVIRA I (cf. [Krause et al. 1997]). Hierbei wurde zusammen mit Verbandsmitarbeitern und Marktforschern in Unternehmungen evaluiert, ob eine Integration von Texten in das für die Recherche von Zeitreihen zu statistischen Erhebungen optimierte System ELVIRA I sinnvoll ist. Interviews und die Analyse von Anfragen von Firmen an die Verbände ergaben, daß textuelle Daten nicht nur explizit nachgefragt werden, sondern oft auch durch die Informationsvermittler nachgewiesen werden, wenn zu einem Thema keine Zeitreihen vorliegen oder die Texte zusätzliche Informationen enthalten. Aufbauend auf dieser Beobachtung wurden einige konkrete Beispiele im Detail analysiert und dienten als Grundlage für die allgemeine Modellbildung (cf. [Krause et al. 1997] und [Mandl et al. 1998]).

2 Das Verbandsinformationssystem ELVIRA

ELVIRA ist ein Client/Server-System zur Recherche und Weiterverarbeitung von Zeitreihen, die überwiegend vom Statistischen Bundesamt, den Industrieverbänden und Wirtschaftsforschungsinstituten erhoben und veröffentlicht werden. Das ursprünglich für den Zentralverband der Elektrotechnik- und Elektronikindustrie (ZVEI, Frankfurt) entwickelte System bieten inzwischen auch der Hauptverband der Deutschen Bauindustrie (HVB, Berlin) und der Verband

Deutscher Maschinen- und Anlagenbau (VDMA, Frankfurt) ihren Mitgliedsunternehmen als Online-System an. Insgesamt ist ELVIRA I an etwa 200 Arbeitsplätzen in Unternehmen und Verbänden installiert und wird von Marktforschern bzw. Informationsvermittlern benutzt (cf. [Scheinost et al. 1998]).

Der Schwerpunkt bei der Entwicklung von ELVIRA lag auf der software-ergonomischen Gestaltung der Benutzungsoberfläche. Der Benutzer soll möglichst schnell und ohne die interne Struktur der Datenbestände zu kennen die für ihn relevanten Daten aus der großen Gesamtmenge (teilweise weit über 5 Mio. Zeitreihen pro Verband) recherchieren. Zu diesem Zweck liegen die Zeitreihen in deskribierter Form vor, wobei jede Zeitreihe durch ein Deskriptorentripel beschrieben wird. Die Deskriptoren jedes Datenbestands (z.B. Produktion, Auftragseingang usw.) lassen sich in drei Facetten gliedern (Thema, Branche und Land) und hierarchisch ordnen. Das Suchobjekt für Zeitreihen (Abb. 1) stellt diese Deskriptorlisten dar, wobei die Liste mit der Facette *Land* aus Platzgründen verdeckt ist.

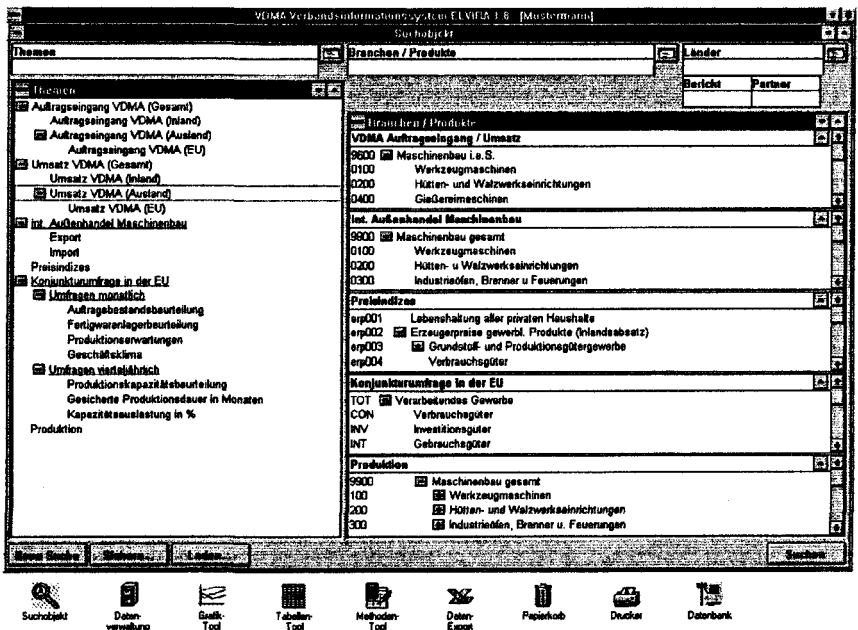


Abb. 1: ELVIRA Recherchebildschirm

Die Vorlageleistung des Systems wurde dahingehend optimiert, daß dem Benutzer zu jedem Zeitpunkt der Anfrageformulierung die Auswahl aus Deskriptorenlisten ermöglicht wird, da die Deskriptoren aufgrund ihrer Komplexität i.d.R. nicht fehlerfrei eingegeben werden können (z.B. „Gleichstrommotoren mit einer Leistung von 0,075 KW bis 0,75 KW“).

Zusätzlich erlaubt das System die Selektion innerhalb der drei Facetten in beliebiger Reihenfolge, so daß der Anwender Fragestellungen wie „Die Produktion in 1000 DM von Produkt X in Deutschland“ oder „Zu Produkt X die Produktion in 1000 DM in Deutschland“ direkt formulieren kann und nicht an eine hierarchische Abfolge gebunden ist.

Um zu gewährleisten, daß der Benutzer in allen drei Facetten die Deskriptoren jeweils eines Bestandes kombiniert, paßt sich die Oberfläche dynamisch an und zeigt bzw. verdeckt die einzelnen Deskriptorenlisten (Abb. 2). Dieser Vorgang kann in jeder Facette ausgelöst werden und erlaubt die gleichzeitige Recherche in beliebig vielen Datenbeständen. Die selektierten Begriffe werden in einer Zustandsanzeige im oberen Bereich des Fensters angezeigt, wobei sich die Höhe der Zustandsanzeige und der Vorlagelisten gegenseitig beeinflußt.

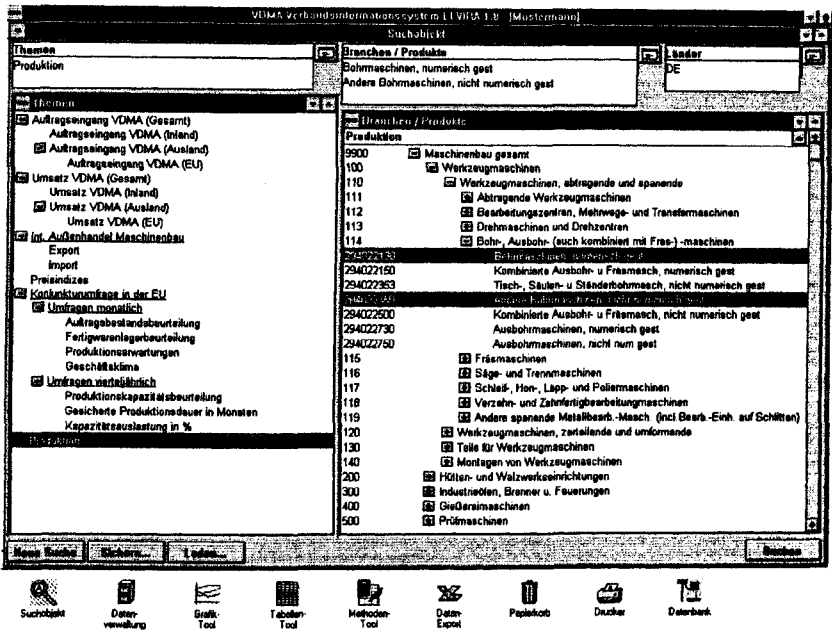


Abb. 2: Recherreformulierung bei der Zeitreihensuche

Ziel von ELVIRA II ist die Integration der Textrecherche in ELVIRA und die Gestaltung der Text-Fakten-Integration. Durch thesaurusbasierte und statistische Verfahren wird auf der Benutzungsoberfläche zum einen die Formulierung von Anfragen über heterogene Datentypen (z.B. Zeitreihen, Texte und Messestermine) ermöglicht und zum anderen die automatische Transformation zwischen unterschiedlichen Indexierungsvokabularien unterstützt.

3 Konzept für die Text-Fakten-Integration in ELVIRA

Die Modellbildung zur Text-Fakten-Integration in ELVIRA II stützt sich auf umfangreiche Interviews mit Nutzern von ELVIRA I und Mitarbeitern der Industrieverbände. Zusätzlich wurden über 100 Anfragen von Firmen an die Verbände ausgewertet, um sowohl konkrete Informationsbedürfnisse als auch Problemlösestrategien bei deren Beantwortung zu ermitteln. Die Analyse relevanter Text- und Datenbestände vervollständigte die Datengrundlage.

Als Ergebnis konnten drei Stufen der Text-Fakten-Integration formuliert werden (cf. [Krause et al. 1997]), die sowohl Benutzern mit spezifischen Vorerfahrungen aus ELVIRA I oder der Nutzung anderen Hosts gerecht werden als auch neue Konzepte der Informationssuche integrieren. Der im allgemeinen Modell postulierten kombinierten Suche entsprechen bei ELVIRA keine Informationsbedürfnisse. Die erste Stufe bilden die Reinformen der Recherche.

A. Text- und Faktensuche in Reinform

Auch in integrierten Informationssystemen treten immer Informationsbedürfnisse auf, die auf eine Modalität beschränkt sind, bei denen der Anwender als Ergebnis also nur Texte oder Fakten erhalten will. Dazu müssen Zugänge bereitstehen, die für diese Aufgabe optimiert sind.

B. Transformationen zwischen den Suchtypen

Transformationen unterstützen den Benutzer dabei, eine für eine Modalität formulierte Anfrage (z.B. nach Zeitreihen) in eine andere Modalität (z.B. nach erklärenden Texten) umzusetzen. Das System transformiert zwischen formalen Unterschieden in den Anfragemöglichkeiten und zwischen dem unterschiedlichen Indexierungsvokabular. Die Transformationen laufen hierbei teilweise transparent im Hintergrund ab oder geben dem Benutzer zusätzliche Möglichkeiten, Parameter der gewünschten zusätzlichen Modalitäten explizit zu setzen. Die spezialisierten Zugänge unterstützen die Anfrageformulierung durch eine dynamisch adaptierte Oberfläche, die z.B. im Zeitreihenzugang boole'sche Operatoren für die Suche nach Texten zuläßt, sowie kontextsensitiv angepaßte Eingabehilfen, wie z.B. Thesauri und Cross-Konkordanzen.

C. Modalitätsunabhängige Suche

Die "echte" Integration geht nur vom Informationsbedürfnis des Benutzers aus, ohne bei der Anfrageformulierung auf eine bestimmte Modalität des Ergebnisses abzielen und durch die Oberflächengestaltung bereits einen Prototypen der Ergebnismenge im mentalen Modell des Benutzers zu implizieren.

Die Abb. 3 zeigt das Modell der Text-Fakten-Integration in ELVIRA II im Überblick und stellt neben den Reinformen (Stufe A) die möglichen Transformationen (Stufe B) schematisch dar. Die Transformationen auf Stufe B erfolgen hauptsächlich durch die automatische Übersetzung von Anfragen, wie in den

folgenden Abschnitten erläutert. Teilweise werden sie aber auch durch feste, intellektuell erstellte Verknüpfungen zwischen inhaltlich zusammengehörenden Text- und Fakten-Dokumenten realisiert, die Experten z.B. bei der Erstellung eines Textes zu den darin verwendeten Fakten definieren.

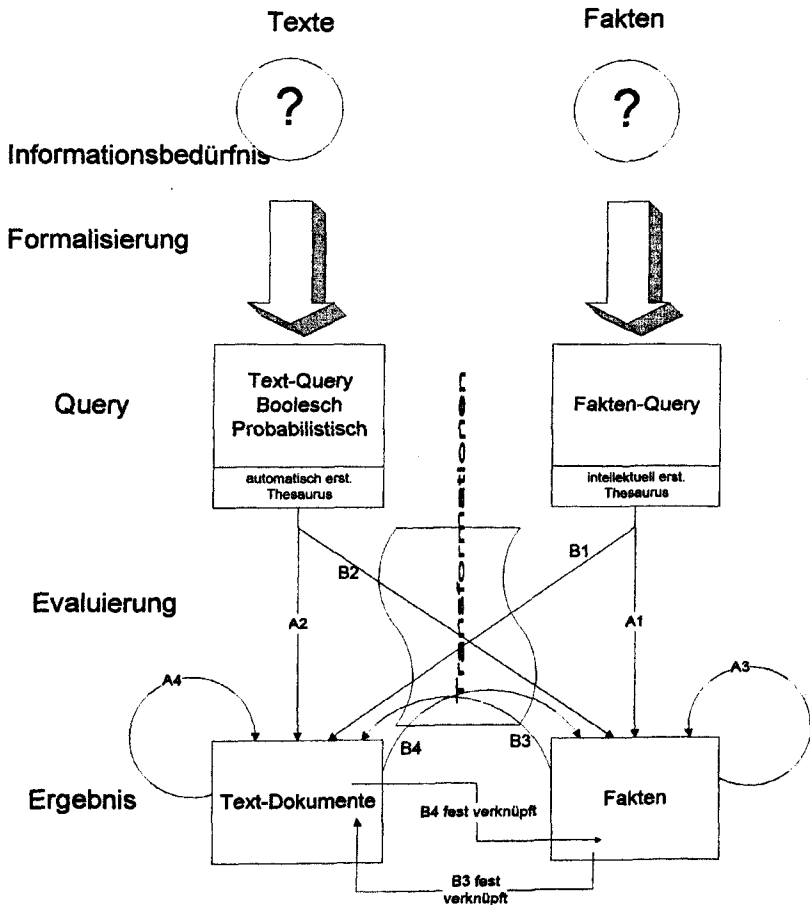


Abb. 3: Reinformen und Transformationen

3.1 Text- und Faktensuche in Reinform

Die Prozesse bei den Reinformen der Recherche lassen sich formal wie folgt beschreiben (vgl. Abb. 3):

- A1: Fakten-Anfrage führt zu Fakten-Ergebnis
- A2: Text-Anfrage führt zu Text-Ergebnis

Die bisherige ELVIRA-Faktenanfrage entspricht dem Typ A1 und die in Entwicklung befindliche Textretrieval-Komponente dem Typ A2.

In der Praxis laufen informationelle Prozesse iterativ ab, d.h. ein Anwender bewertet meist das aktuelle Ergebnis und formuliert aufgrund seiner Wertung die ursprüngliche Anfrage um. Erst nach evtl. mehreren Zyklen erhält er ein befriedigendes Ergebnis, das im Falle von ELVIRA nicht bereits dem gewünschten Endergebnis entsprechen muß (oft schließen sich hier noch zusätzliche Verarbeitungsschritte an).

Eine spezifische Art des iterativen Retrievals ist das im Textretrieval verbreitete Relevance Feedback. Dabei markiert der Anwender in den Ergebnismengen die für ihn relevanten Dokumente. Das System nutzt diese Informationen zur Formulierung der neuen Anfrage. Der Benutzer gelangt mit dem aktuellen Ergebnis - oder Teilen davon - direkt zu einem neuen Ergebnis. Diese Technik ist auch im Faktenretrieval von ELVIRA sinnvoll, da sich durch die in Kapitel 2 erläuterte Verschlagwortung von Zeitreihen die Verfahren aus dem Textretrieval direkt übernehmen lassen. Somit ergeben sich zwei weitere Suchtypen (vgl. Abb. 3):

- A3: Fakten-Ergebnis führt zu Fakten-Ergebnis
- A4: Text-Ergebnis führt zu Text-Ergebnis

3.2 Transformationen zwischen den Suchtypen

Automatische Transformationen können Anfragen einer Modalität in andere Modalitäten übersetzen und den Benutzer entlasten. Mit einer Textanfrage ist so z.B. die gleichzeitige oder iterative Suche nach Texten und Zeitreihen möglich. Die Aufgabe der Transformationsfunktionen ist es hierbei, die Eigenheiten der Datenbestände möglichst gut aufeinander abzubilden. Strukturelle und semantische Unterschiede verhindern oft eine direkte Überführung, wobei sich die Probleme i.d.R. erst nach eingehender Analyse darstellen und zumindest für den gelegentlichen Benutzer nicht sofort erkennbar sind. Neben strukturellen Unterschieden im Aufbau verhindern vor allem unterschiedliche Indexierungsvokabulare den direkten Transfer. So weicht z.B. die Wortwahl in amtlichen Nomenklaturen zu Zeitreihendaten vom Sprachgebrauch in der Praxis ab. Dies hat zur Folge, daß z.B. mit dem Suchbegriff „Waschmaschine“ zwar Texte gefunden werden können, Faktendaten aber aufgrund der Verwendung des Begriffs „Waschvollautomat“ nicht nachgewiesen werden.

Beim Transfer von Fakten zu Texten ergeben sich ähnlich gelagerte Probleme, wenn Attributausprägungen von Fakten(-beschreibungen) bei Textdaten im Volltext und zusätzlich als Schlagwort erscheinen können. Auch hier ist nach eingehender Analyse eine dedizierte Transformationsfunktion zu spezifizieren, welche die primäre Suchintention des Benutzers optimal widerspiegelt.

Unter Einbeziehung der iterativen Fälle ergeben sich aus Nutzersicht somit folgende Suchprozesse (vgl. Abb. 3):

- B1: Fakten-Anfrage führt zu (Fakten- und) Text-Ergebnis
- B2: Text-Anfrage führt zu (Text- und) Fakten-Ergebnis
- B3: Fakten-Ergebnis führt zu Text-Ergebnis
- B4: Text- Ergebnis führt zu Fakten-Ergebnis

3.3 Modalitätsunabhängige Suche

Auf der Stufe C der Integration fokussiert der Anwender nur noch auf sein Informationsbedürfnis und nicht auf die Modalität der Ergebnisdokumente (vgl. Abb. 4):

- C: Informationsbedürfnis führt zu Text-Ergebnis + Fakten-Ergebnis

Dieses Suchbedürfnis kann der Anwender wohl nicht in einem für eine bestimmte Modalität optimierten Zugang formulieren. Vielmehr müßte ein Werkzeug für eine abstrakte Suchanfrage entwickelt werden, das nach Möglichkeit keinen spezifischen Ergebnistyp bereits bei der Anfrageformulierung impliziert.

Erste empirischen Untersuchungen haben gezeigt, daß solche Bedürfnisse im Anwendungsfall von ELVIRA vorkommen. Ob Benutzer die Suche auch so konzeptualisieren, ist dagegen noch nicht klar. Es ist möglich, daß der Suchformulierung bereits ein Prototyp des Ergebnisses zugrunde liegt und damit das abstrakte Problem auf der Ebene der Suchformulierung schon keine Entsprechung mehr hat. Hierzu ein Beispiel:

Ein Benutzer will "alles zu einem Land" wissen. Er erhält in diesem Fall von den Informationsvermittlern der Verbände ein Paket mit verschiedensten Informationen. Dabei schränken die Informationsvermittler die Anfrage mit oder ohne Rücksprache meist auf das Produktspektrum des anfragenden Unternehmens ein.

Ob diese Art der Integration für ELVIRA II sinnvoll ist, können erst weitergehende empirische Untersuchungen zeigen. In einer ersten Phase sollen hierzu entsprechende Tests bei ELVIRA-Nutzern durchgeführt werden.

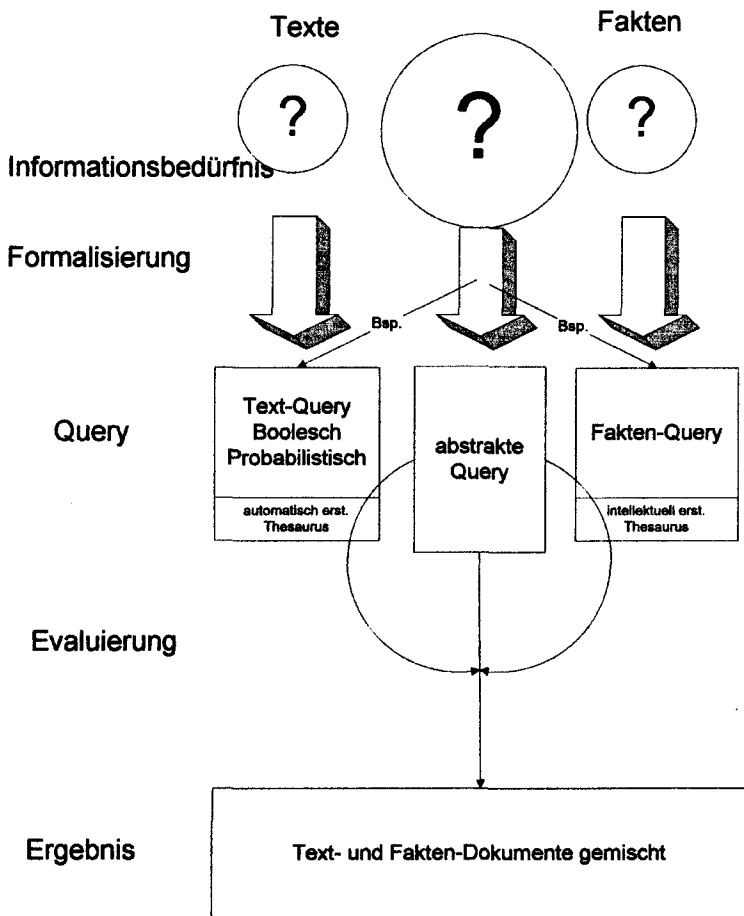


Abb. 4: Modalitätsunabhängige Suche

3.4 Zusammenfassung und weiteres Vorgehen

Empirische Untersuchungen im Anwendungsfall ELIVRA haben gezeigt, daß von den Integrationsstufen der kombinierten Text-Fakten-Recherche besonders die Transformationen auf der Stufe B für die Benutzer sinnvoll sind. Mit der Modellbildung gingen umfangreiche Analysen unterschiedlichster Datenarten und -quellen einher. Wie zu erwarten, zeigte sich, daß sowohl der formale Aufbau der Datenbestände als auch die Qualität der Erschließung erhebliche Unterschiede aufweist. Die Spannbreite reicht von nach formalen Regeln erschlossenen und konsistent formatierten Daten (z.B. Zeitreihen, Messetermine, einzelne Textbestände) bis zu sehr grob bzw. ungegliederten Daten-

typen (z.B. Texte aus News-Tickern). Einige Felder aus Faktendaten finden ihre Entsprechung in bibliographischen Angaben oder Schlagwörtern von Textbeständen, während andere Textbestände nur über automatische Indexierung erschlossen werden können. Dies bedingt den Einsatz unterschiedlichster Verfahren zur Transformation von Anfragen. Der Schwerpunkt der aktuellen Entwicklung zur Realisierung von Transformationen liegt auf den folgenden Bereichen:

- Über ein umfassendes Metamodell werden die heterogenen Datenbestände beschrieben und integriert, wobei neben dem formalen Aufbau auch Angaben über anwendbare Transformationsalgorithmen und verfügbare Hilfsmittel (z.B. Klassifikationen, Thesauri oder Konkordanzen) abgelegt sind.
- Thesaurusbasierte Verfahren, die dem Benutzer den Zugang über seine eigene Begrifflichkeiten ermöglichen. Hierzu befindet sich ein Thesaurus im Aufbau, der zunächst einzelne Teilbereiche aus dem Spektrum der beteiligten Verbände abdeckt und durch eine Konkordanz zu unterschiedlichen Datenbeständen vervollständigt wird.
- Vage Verfahren auf der Basis von Kookurenzen und neuronalen Netzen, die mit möglichst geringem Pflegeaufwand von Seiten der Verbände Suchanfragen zwischen dem unterschiedlichen Indexierungsvokabular der Datenbestände transformieren können (cf. [Mandl 1998]). Diese Ansätze extrahieren, die ihr Wissen in Form von Kookurenzen aus Doppelkorpora, die mit zwei Verfahren indexiert wurden. Ein erfolgversprechender statistischer Ansatz von [Sheridan/Ballerini 1996] wurde bereits für Multilinguales Retrieval eingesetzt und wird für ELVIRA adaptiert.

4 Textrecherche mit integrierter Faktensuche

Dieser Abschnitt veranschaulicht den Ablauf einer integrierten Suche anhand einer skizzenhaften Visualisierung der Interaktionsschritte. Das Szenario stammt aus [Mandl et. al. 1998], wo Beispiele für alle Typen der Integration gesammelt wurden. Sie wurden zusammen mit Verbandsmitarbeitern auf Plausibilität und Praxisrelevanz geprüft.

Die in Kapitel 3 aufgeführte Reinform zur Textrecherche wurde mittlerweile prototypisch realisiert und in ELVIRA II integriert (cf. [Krause/Schaefer 1998]). Im gewählten Beispiel sucht der Benutzer primär nach Texten und öffnet dementsprechend das Text-Suchobjekt. Daneben möchte er aber auch Zeitreihen erhalten, die relevant bzgl. seiner Problemstellung sind (Auswahl in der Liste Dokumenttyp; vgl. Abb. 5, oben rechts). Der Anwender aus dem Bereich der Elektroindustrie interessiert sich für den Markt für Waschmaschinen in Ungarn.

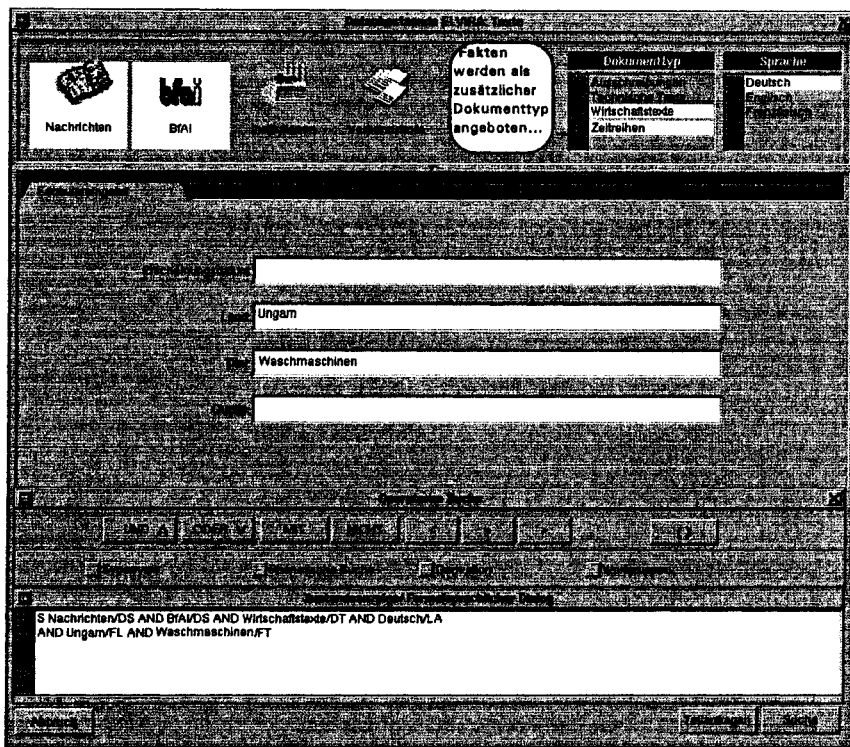


Abb. 5: Text-Suchobjekt (prototypischer Entwurf)

Beim Suchprozeß wird die Text-Anfrage intern in eine Zeitreihen-Anfrage transformiert. Die in Abb. 6 dargestellte Zustandsanzeige der Zeitreihensuche (siehe auch Abb. 2) macht dies deutlich.

Themen	Branchen / Produkte	Länder
Import	Haushalts-Waschgeräte	Bereich
Export	Waschvollautomaten mit Fassungsvermögen bis 6 kg (bis '93)	Partner
	Waschvollautomaten mit Fassungsvermögen bis 6 kg Frontlader (ab '94)	D
	Waschvollautomaten mit Fassungsvermögen bis 6 kg Toplader (ab '94)	Ung
	Waschvollautomaten mit Fassungsvermögen über 6 kg - 10 kg	
	Waschmaschinen mit eingebauter Wäschescneider	
	Andere Waschmaschinen	
	Teile für Waschvollautomaten (ab '95)	

Abb. 6: Interne Zeitreihen-Anfrage

Die Leistung der Transformation besteht in diesem Fall in der Ausweitung von *Waschmaschinen* auf den Begriff *Waschvollautomaten*, der in der Nomenklatur verwendet wird. Dies kann durch eine Synonym-Relation im Thesaurus erreicht werden. Liegt kein Thesaurus vor, können aus dem IR bekannte Verfahren zur Termexpansion eingesetzt werden. Diese basieren meist auf Kookurrenzen und gehen davon aus, daß Begriffe, die häufig gemeinsam in Dokumenten vorkommen, ähnliche Bedeutung haben.

[illegible]

Das Design der Benutzungsoberfläche für die Integration ist neben der Realisierung von Transformationen der Schwerpunkt der momentanen Entwicklung von ELVIRA. Das obige Beispiel zeigt die prototypische Implementierung der Transformationen von Texten zu Fakten. Eine ähnliche Funktionalität wird in das Zeitreihenrecherche-Werkzeug integriert. Diese Prototypen werden mit Benutzern getestet, wobei auch die Qualität der Transformationen untersucht wird.

[Amstutz et.al. 1991]

[Fuhr 1992]

- Fuhr, Norbert: *Integration of Probabilistic Fact and Text Retrieval*. In: Belkin, Nicholas; Ingwersen, Peter; Pejtersen, Annelise (Hrsg.) (1992): SIGIR'92. Copenhagen, Denmark. June 21-24 1992. S. 211-222.
- [Gu et al. 1993]
Gu, J.; Thiel, U.; Zhao, J.: *Efficient Retrieval of Complex Objects: Query Processing in a Hybrid DB and IR System*. In: Knorz, Gerhard; Krause, Jürgen; Womser-Hacker, Christa (Hrsg.) (1993): Information Retrieval '93. Von der Modellierung zur Anwendung. Proc. der 1.Tagung Information Retrieval. Konstanz. S. 67-81.
- [Krause et al. 1997]
Krause, Jürgen; Mandl, Thomas; Stempfhuber, Maximilian: *Text-Fakten-Integration in ELVIRA*. ELVIRA-Arbeitsbericht 12, IZ Sozialwissenschaften, Bonn, 1997.
- [Krause et al. 1998]
Krause, Jürgen; Mandl, Thomas; Schaefer, André; Stempfhuber, Maximilian: *Stand des Projektes ELVIRA II*. ELVIRA-Arbeitsbericht 19, IZ Sozialwissenschaften, Bonn, 1998.
- [Krause/Schaefer 1998]
Krause, Jürgen; Schaefer, André: *Textrecherche-Oberfläche für ELVIRA II*. ELVIRA-Arbeitsbericht 16, IZ Sozialwissenschaften, Bonn, 1998.
- [Mandl 1998]
Mandl, Thomas: *Vague Transformations in Information Retrieval*. In this volume.
- [Mandl et al. 1998]
Mandl, Thomas; Schaefer, André; Stempfhuber, Maximilian: *Exemplarische Transformationen für die Text-Fakten-Integration*. ELVIRA-Arbeitsbericht 15, IZ Sozialwissenschaften, Bonn, 1998.
- [Marx et al. 1995]
Marx, Jutta; Mutschke, Peter; Schommier, Marcus: *Möglichkeiten der intelligenten Integration heterogener Datenbestände: das Projekt GESINE*. IZ-Arbeitsbericht Nr. 2, IZ Sozialwissenschaften, Bonn.
- [Scheinost et al. 1998]
Scheinost, Ulrich; Haas, Hansjörg; Krause, Jürgen; Lindlbauer, Jürg (Hrsg.): *Marktanalyse und Marktprognose. Das ZVEI Verbandsinformationssystem ELVIRA*. IZ Sozialwissenschaften, Bonn, 1998.
- [Sheridan/Ballerini 1996]
Sheridan, Páraic, Ballerini, Jean Paul: *Experiments in Multilingual Information Retrieval using the SPIDER System*. In: Frei et al. 1996. p. 58-65.
- [Womser-Hacker 1997]
Womser-Hacker, Christa: *Das MIMOR-Modell. Mehrfachindexierung zur dynamischen Methoden-Objekt-Relationierung im Information Retrieval*. Habilitationsschrift. Universität Regensburg. Informationswissenschaft.

Qualitätsbewertung von Information-Retrieval-Systemen

Sind Synergieeffekte durch die Koordination verschiedener Bewertungsmethoden möglich?

Osman Bayraktar, Christa Womser-Hacker

Osman Bayraktar

IDS Prof. Scheer GmbH; Altenkesseler Str. 17; 66115 Saarbrücken
o.bayraktar@ids-scheer.de

Christa Womser-Hacker

Universität Konstanz; Informationswissenschaft; Postfach D 78; D-78457 Konstanz
cwh@inf-wiss.uni-konstanz.de

Inhalt

- 1 Einleitung
- 2 Zielrichtungen der Qualitätsbewertung
 - 2.1 Auswahl eines IRS
 - 2.2 Traditionelle IR-Evaluierung
- 3 Kombination: Ein Blick in die Zukunft
- 4 Fazit

Zusammenfassung

Ziel dieses Beitrags ist die Gegenüberstellung verschiedener Bewertungsmethoden und Zweckbestimmungen im Bereich von Information-Retrieval-Systemen. Es wird aufgezeigt, welche Ansätze vorliegen und hinterfragt, wie sie einander angepaßt werden können, um synergetische Effekte zu erzielen.

1 Einleitung

Zunehmender Wettbewerbsdruck und die Globalisierung von Märkten haben zur Folge, daß die Kundenorientierung im Produktionsmanagement, und hier vor allem im Qualitätsmanagement, stärker als je zuvor Berücksichtigung findet. Die Aufgaben unterliegen einem permanenten Wechsel; ihre Erfüllung hängt immer mehr von der Güte und Qualität der bereitgestellten Informationen ab. Der Einsatz von Informationstechnologien und das Management von Wissen haben sich daher für die Unternehmensführung zu einer strategischen Aufgabe entwickelt. Dabei liegt es am Informationsmanager, die strategische Rolle der Informationsfunktion und die inner- und außerbetrieblichen Bedingungen für die Umsetzung des Leistungspotentials der Informationsinfrastruktur zu analysieren und zu bestimmen. Nach einer Umfrage der Gartner Group liegen 90% alles Wissens in textueller Form vor. Deshalb werden Information-Retrieval-Systeme (IRS) zunehmend neben anderen Typen von Informationssystemen als integrale Bestandteile der Informationsinfrastruktur moderner Betriebswirtschaften angesehen, welche die nötigen Informationen (Informationsnachfrage) für die Bewältigung der Aufgaben wirtschaftlich auf der operativen, administrativen und strategischen Ebene bereitstellen. Dabei wird beim begrifflichen Verständnis von IRS über die traditionelle Sichtweise des Text- und Dokumentenretrieval hinausgegangen und auf Systeme erweitert, welche Information in verschiedenen Medien und Sprachen integrieren und auch Mischformen (wie Texte, Bilder, Fakten) zulassen.

Seit einigen Jahren existieren IRS mit komplexer Funktionalität (z.B. gewichtete Indexierung, Ranking, Relevance Feedback etc.) nicht nur als Prototypen im Bereich der Forschung. Einige Systeme haben einen Stand der Reife erreicht, so daß fertige Produktlösungen auch im kommerziellen Bereich zur Verfügung stehen. Durch diese Situation auf dem Informationsmarkt und durch die Neuorientierung vieler Organisationen und Unternehmen, der Informationsflut Herr zu werden, ergibt sich auch im Bereich der (vergleichenden) Qualitätsbewertung dieser Produkte eine zu überdenkende Situation im Hinblick auf die einzusetzenden Methoden.

2 Zielrichtungen der Qualitätsbewertung

Bei der Qualitätsbewertung von IRS oder IRS-Komponenten stehen sich unterschiedliche Zielrichtungen gegenüber, die wiederum in Kombination mit unterschiedlichen Methoden auftreten. Das Spektrum reicht hier von kontrollierten Experimenten, Untersuchungen, empirischen Langzeitstudien, Rapid-Prototyping-Verfahren und Simulationstests bis hin zu betriebswirtschaftlich orientierten Vorgehensweisen wie Benchmarking und der Analyse von Kosten-Nutzen-Faktoren oder (kritischer) Erfolgsfaktoren.

Die Zweckbestimmungen dieser Ansätze reichen von Anschaffungsentscheidungen, Funktionalitätsprüfungen, Zertifizierungsbestrebungen z.B. im Rahmen von Vermittlungsdiensten bis hin zu Initiativen genereller Natur, bei welchen das wissenschaftliche Interesse an der Funktionsweise des Retrievalprozesses do-

minierte. All diese Vorgehensweisen sind derzeit nur schwierig übertragbar und für andere Zwecke zu verwenden, da sie äußerst unterschiedlichen Bewertungsphilosophien folgen.

Exemplarisch werden hier zwei relativ weit auseinander liegende Zielrichtungen (die Anschaffungsentscheidung eines IRS und die TREC-Studie) gegenübergestellt. Anschließend soll geprüft werden, an welchen Stellen Veränderungen möglich sind, die synergetische Effekte für eine Mehrfachnutzung und -interpretation erwarten lassen.

2.1 Auswahl eines IRS

Bei der Neuanschaffung von IRS stellt sich die Frage nach der Vorgehensweise, mit der prognostiziert werden kann, wie sich ein System im gegebenen Kontext verhalten wird. Idealerweise besteht die Möglichkeit eines Retrievaltests, welcher die auf dem Markt vorhandenen Systeme mit den tatsächlichen Gegebenheiten des intendierten Einsatzes in Beziehung setzt. Dabei wird auch die vorliegende Infrastruktur sowohl in bezug auf Soft- und Hardwareausstattung (mit Blick auf die Kompatibilität) als auch die intendierte Nutzungssituation berücksichtigt. Diese Strategie wurde mit der PADOK-Studie zur Ermittlung eines Indexierungsverfahrens für das Deutsche Patentinformationssystem verfolgt (cf. KRAUSE/WOMSER-HACKER 1990). Aufgrund der Aufwendigkeit und Kostspieligkeit derartiger Tests kommen solche Verfahren selten zum Einsatz, obwohl sie optimale Ergebnisse erwarten lassen.

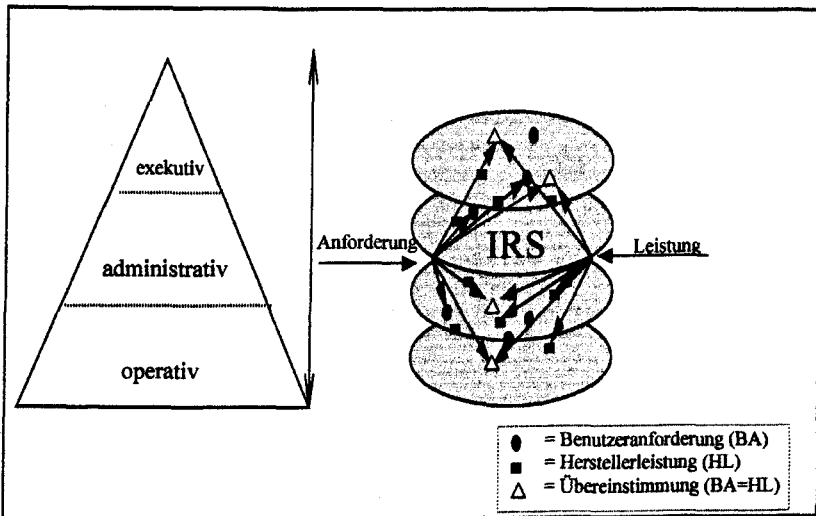


Abb. 1: Benutzersicht/-anforderung vs. Herstellersicht/-leistung

Abb.1 stellt die Situation dar, deren Ziel es ist, ein IRS auszuwählen, bei welchem sich möglichst viele Anforderungen von Seiten der Benutzer mit Herstellerleistungen decken. Dabei können die „Benutzer“ den verschiedenen Ebenen einer Organisation angehören oder was den häufigeren Fall darstellt, die entscheidende Gruppe nicht identisch mit der das IRS hauptsächlich nutzenden Gruppe sein. Letzteres erfordert einen ebenenübergreifenden Kommunikationsprozeß. Als methodisches Instrumentarium bietet sich eine Erfolgsfaktorenanalyse an um zu klären, welche IRS-Eigenschaften von den betroffenen Mitgliedern der unterschiedlichen organisatorischen Ebenen als relevant eingestuft werden. Für diese Eigenschaften werden Qualitätsaussagen benötigt, mit dem Ziel, sie miteinander zu kombinieren, um eine die Prioritäten berücksichtigende Gesamtbewertung zu erhalten.

2.2 Traditionelle IR-Evaluierung

Dem gegenüber stehen traditionelle Evaluierungskonzepte des IR, die meist nur in sehr geringem Maße an der Berücksichtigung von Benutzerinteressen orientiert sind. Im Vordergrund steht die Effektivität der Retrievalmaschinen, d.h. des Retrievalkerns, der für den Match zwischen Informationsobjekten und Anfragen verantwortlich ist (z.B. Vector Space Model vs. Probabilistisches Retrievalmodell) und das Ziel verfolgt, relevante Objekte zu selektieren und gleichzeitig irrelevante zurückzuhalten.

Im Information Retrieval wird die Entwicklung der Funktionalität seit jeher begleitet von Evaluierungsaktivitäten (cf. z.B. die Arbeiten zum SMART-System in SALTON 1971). Zu diesem Zwecke hat sich ein komplexer Methodenkanon herausgebildet, der in erster Linie auf statistischen Verfahren basiert, die auf den Information-Retrieval-Bereich übertragen wurden (Basismaße recall und precision, single-number-measures wie das E-Maß nach VAN RIJSBERGEN etc.). Neben der Forschung im Bereich der mathematischen Meßtheorie (cf. BOLLMANN 1977 und 1981, BOLLMANN/CHERNIAVSKY 1981) wurden Benutzeraspekte nur sehr rudimentär in die Maßzahlen integriert. Ein Beispiel ist die Formalisierung von sog. Benutzerstandpunkten z.B. im normalisierten recall, welche die Position von Ergebnis-Objekten in Ranking-Listen miteinbeziehen. Als Benutzerstandpunkt gilt z.B. eine spezifische Benutzerverhaltensweise, die Auswertung der Ranking-Liste nach drei aufeinander folgenden irrelevanten Dokumenten abzubrechen. Für die Berechnung der Maße werden nur die Objekte bis zum festgelegten Cut-off einbezogen. Daß eine derartige Formalisierung aufgrund ihrer Pauschalität das reale Benutzerverhalten nur sehr bedingt abbilden kann, versteht sich von selbst.

Als derzeit größte Evaluierungsinitiative und als aktueller Endpunkt der langen Evaluierungstradition im IR sind die Text-Retrieval-Konferenzen (TREC) anzusehen (cf. HARMAN 1995 und 1998). Sie sollen dargestellt und in diesem Rahmen diskutiert werden (auch in bezug auf ihre Tauglichkeit für Anschaffungsentscheidungen).

TREC hat sich in den USA seit Anfang der neunziger Jahre als Initiative etabliert, die von staatlicher Seite (ARPA Advanced Projects Agency) am NIST (National Institute of Standards and Technology) gefördert wird. Auf der Basis großer Dokumentkollektionen (Dokumente, Topics und Relevanzurteile) wurden beispielsweise in der sechsten TREC-Runde 58 Retrievalsysteme verglichen (cf. VOORHEES/HARMAN 1998). Die teilnehmenden Systeme waren Forschungssysteme aus Universitäten und wissenschaftlichen Industrieabteilungen, aber auch kommerzielle Systeme nahmen teil. Von allen Systemen wurden umfangreiche Angaben über die notwendigen Ressourcen (man and machine power) gefordert. Benutzer spielen bisher bei TREC noch keine Rolle, da die Erarbeitung der Anfragen aus den Topics von den jeweiligen Entwicklern geleistet wird. Diese versuchen, auf der Basis ihrer optimalen Kenntnis des jeweiligen Systems dessen Wirkung so positiv wie möglich zu steuern. Benutzer, die ihr System oft als Black Box wahrnehmen, reagieren folglich auf völlig andere Art und Weise. Da im Rahmen von TREC auch die Evaluierungsmethode auf die Probe gestellt wird, erfolgt über die Phasen hinweg eine gezielte Variation der Testgrundlagen (z.B. Verwendung starkt elaborierter und wenig elaborierter Topics). Das aus TREC zur Verfügung stehende Material ist äußerst umfangreich, jedoch noch sehr wenig analysiert. Es beinhaltet neben den Testgrundlagen, den Fragebögen zu den einzelnen Systemen einen Leistungsvergleich der Systeme auf der Basis statistischer recall-precision-Werte. Die wichtigsten Maßzahlen bilden die Ergebnisse (inkl. ihrer graphischen Darstellung) der sog. frozen ranks method. Hier werden an vorher festgelegten Rängen Einschnitte vorgenommen und punktuelle precision-Werte berechnet, die anschließend einer Mittelung unterzogen werden. Als sehr aufschlußreiche, aber zeitaufwendige Informationsquelle sind in den TREC-Proceedings Beiträge der einzelnen Teilnehmergruppen enthalten, welche die jeweiligen Vorgehensweisen ihrer Gruppe im Detail beschreiben.

TREC ist auf Ranking-Systeme spezialisiert, d.h. Boolesches Retrieval nimmt am Test nicht teil. Insofern steht hier die Fähigkeit der Systeme im Vordergrund, möglichst viel relevantes Material auf guten Rangplätzen zu positionieren. Hier spielen neben der Gewichtungsfunktion in erster Linie Relevance-Feedback-Ansätze und Anfrageoptimierungstechniken eine entscheidende Rolle. In der Praxis liegt meist die notwendige gewichtete Indexierung nicht vor.

TREC ergänzt die klassische Evaluierungsform durch die Einbeziehung von Effizienzgesichtspunkten wie Systemressourcen, manueller und maschineller Aufwand, Input/Output-Operationen etc., jedoch sind diese Faktoren, die in einem Fragebogen ermittelt werden, bisher noch in keinster Weise Gegenstand weiterführender Analysen geworden.

Bei TREC wird die Hauptevaluierungslinie durch Nebenstränge (sog. special TREC tracks) ergänzt, die besondere Aspekte berücksichtigen (z.B. Multi- und Crosslingualität, Filtering, Verteilte Datenbanken, Spoken Language, Interaktivität etc.).

Die Interpretation der TREC-Ergebnisse kann unter verschiedenen Perspektiven geschehen. Zum einen werden die Entwicklungsgruppen die Veränderungen der Ergebnisse mit den Veränderungen im Systemdesign über die Phasen hinweg in Beziehung setzen. Wesentlich schwieriger ist es, bestimmte Werte mit isolierten Funktionen zu verbinden, d.h. z.B. eine Verbesserung in der pre-

cision mit einer aufwendigeren linguistischen Analyse zu begründen. Diese Herangehensweise ist möglicherweise äußerst versierten Kennern der TREC-Zusammenhänge möglich.

3 Kombination: Ein Blick in die Zukunft

Die Qualität von Systemen hängt in starkem Maße von den jeweiligen Bedingungen ihres zukünftigen Einsatzes ab. Insofern ist der Kontext und die Rahmenbedingungen von großer Bedeutung für die Auswahl eines geeigneten Systems. Dies ist mit einer vielschichtigen Problematik verbunden.

Zunächst sind die Faktoren zu ermitteln, die für die Bewertung innerhalb einer Organisation oder eines Unternehmens entscheidend sind. Dies erfordert einen iterativen Prozeß, um die Anforderungen exakt zu erarbeiten bzw. herauszudestillieren und miteinander in Beziehung zu setzen. Wichtig ist, daß durch die Beteiligung von Personen mit verschiedenen Aufgaben und den sie einbeziehenden Kommunikationsprozeß sichergestellt wird, daß ein homogenes Verständnis vorliegt, das die Faktoren nicht nur in abstrakter Form anführt, sondern ihre operative Bedeutung miteinbezieht (z.B. Was bedeutet „Benutzerfreundlichkeit“, „Transparenz“, „Selbsterklärungsfähigkeit“ im IR-Kontext?¹). Hinzu kommt die Notwendigkeit, Wissen über die Potentiale aktueller IRS an dieser Stelle einzubringen. Als Methode wurde auf die Erfolgsfaktorenanalyse (EFA) (cf. HEINRICH 1996) zurückgegriffen, mit der es möglich ist, Qualitätsmerkmale von IRS aus der jeweiligen Benutzersicht und vor dem Hintergrund der jeweiligen Arbeitsanforderung zu identifizieren, um somit eine Grundlage für die Systembewertung bereitzustellen.

Die Darstellung der EFA wird durch die Ergebnisse einer Studie untermauert, die im Rahmen des Kurses „*Evaluation interaktiver, multilingualer und multimedialer Informationssysteme*“ an der Universität Konstanz vorgenommen wurde. Die Identifikation der Erfolgsfaktoren erfolgte im Rahmen eines dreiwöchigen Workshop. Teilnehmer² der Arbeitsgruppe waren sieben fortgeschrittene Studenten des Aufbaustudiums der Konstanzer Informationswissenschaft und ein Moderator. Die Teilnehmer verfügten über unterschiedlichen Background; ihre Erstausbildung erstreckte sich von Sozialwissenschaften über Naturwissenschaften bis zu den Geisteswissenschaften, wobei die Mehrzahl bereits über mehrere Jahre Berufspraxis verfügte. Einschränkend ist zu sagen, daß es sich nicht um eine „echte“ geschlossene IRS-Benutzergruppe handelte und daher die Studie als erster Versuch hinsichtlich der Bestimmung der Erfolgsfaktoren aus Benutzersicht zu verstehen ist. Sie stellt eine Orientierungshilfe für eine breit angelegte Studie dar.

Der erste Schritt bestand in einer Brainstorming-Sitzung, in der Qualitätsmerkmale gesammelt wurden. Die Teilnehmer wurden dabei durch die Präsentation verschiedener Beispiele sensibilisiert. Nachdem ein gewisses Vertrauen in die

¹ Auch die ISO-Normen (z.B. 9126, 9241 Teil 10) bleiben hier sehr vage und gehen nicht auf spezifische Ausprägungen von Software-Systemen ein.

² Hier sind Frauen und Männer gemeint.

Methode des Brainstorming vorhanden war, hielt die Gruppe insgesamt 66 in Frage kommende Erfolgsfaktoren fest. Anschließend wurden die aufgelisteten Eigenschaften diskutiert und überarbeitet, wobei folgende Kontrollfragen im Mittelpunkt der Diskussion standen:

- Welche der Eigenschaften sind für die Benutzerzufriedenheit entscheidend?
- Welche Eigenschaften fehlen und müssen ergänzt werden?

Deutlich wurde hier die Notwendigkeit ausführlicher Diskussion, um eindeutige Bedeutungen und Definitionen zu gewährleisten und terminologische Vereinheitlichung zu erzielen. Die Arbeitsgruppe versah jeden Erfolgsfaktor mit einer genauen Beschreibung. Weiteres Ziel des Workshop war, eine Systematisierung und Kategorisierung der Erfolgsfaktoren herbeizuführen. Das Ergebnis dieser Kategorisierung wurde mehrfach überarbeitet.

Das Ergebnis der Studie machte neben den Anforderungen an die Funktionalität die Bedeutung von Benutzerfaktoren, Service und Qualitätsmerkmalen deutlich, wobei in besonderem Maße die Benutzungsschnittstelle und ausführliche Hilfe- und Unterstützungsfunktionen betont wurden. Abbildung 2 zeigt exemplarisch die Ergebnisse auf. Im Detail sind diese im Abschlußbericht (cf. BAYRAKTAR/WOMSER-HACKER 1998) enthalten.

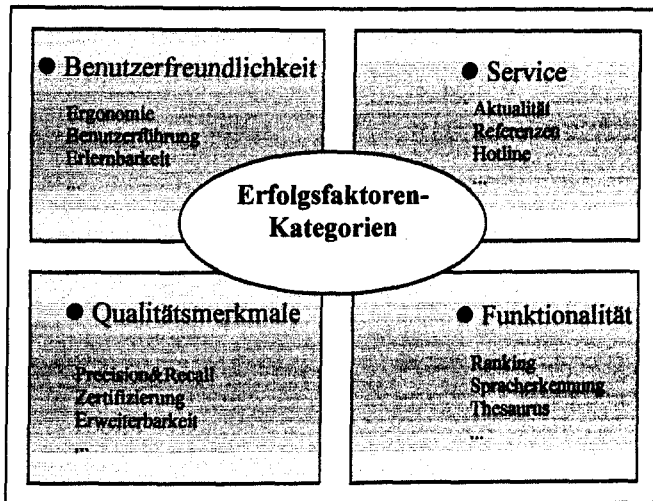


Abb. 2: Kategorisierung der Erfolgsfaktoren

Der nächste Schritt bestand in der Klassifizierung der Qualitätsmerkmale nach ihrer Priorität. Auch hier sollte die Sicht des Benutzers die Grundlage der Bewertung darstellen (Leitfragen: *Was will der Benutzer? Wofür ist er bereit zu warten oder zu bezahlen?*). Die Bewertung der Erfolgsfaktoren erfolgte mit Hilfe von Fragebögen, die einen Prioritätsteil umfaßten.

Die folgende Abbildung zeigt einen Ausschnitt aus den Fragebogen exemplarisch auf:

Bewerten Sie die Erfolgsfaktoren nach ihrer Bedeutung! Welche Bedeutung (Wichtigkeit oder Priorität) haben die Erfolgsfaktoren für das IRS?

Erfolgsfaktor (X)	Irrelevant	evtl. nützlich			wichtig	entscheidend	
	1	2	3	4	5	6	7
Ergonomie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Spracherkennung	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Transparenz der Systemfunktionalität	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Antwortzeit	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abb. 3: Bewertung der Erfolgsfaktoren nach ihrer Bedeutung

Die Ergebnisse wurden anhand eines Portfolios visualisiert, aus dem der Handlungsbedarf deutlich wurde.

Auch generelle Probleme derartiger Analysemethoden konnten in diesem Zusammenhang aufgedeckt werden. Ein besonderes Problem besteht in der Komplexität aktueller IRS, deren Systemfunktionalität sich nur auf einer sehr detaillierten Ebene systematisch darstellen, kommunizieren und vergleichbar machen läßt.

Es stellt sich nun die Frage, inwieweit großangelegte Studien wie TREC einen Beitrag leisten können um zu prüfen, in welchem Maße die erarbeiteten Erfolgsfaktoren erfüllt sind.

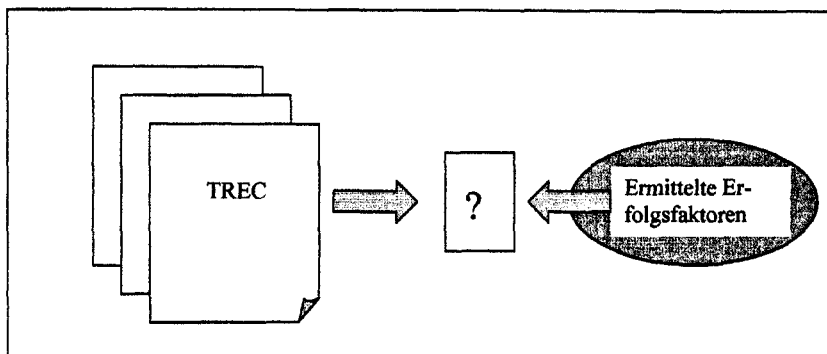


Abb. 4: Kombination von Methoden

Trotz der Fülle des Materials, das durch die sechs bisherigen TREC-Phasen geliefert wurde, fehlen weiterführende Analysen, welche die Ergebnisse auf eine Metaebene stellen, die zweckorientierten Interpretationen zugänglich ist. Noch erfordert dies eine intensive Kenntnis der TREC-Zusammenhänge, um das Material optimal verwerten zu können.

Ein wichtiger Punkt besteht z.B. in der Aufteilung der bei TREC getesteten Systeme, in solche die erworben werden können und somit für einen realen Einsatz überhaupt in Frage kommen können, und reinen Forschungssystemen, die über keine Infrastruktur zur Vermarktung und Pflege verfügen. TREC stellt sich die Aufgabe, auch in diesem Sinne Transferarbeit zwischen Industrie und Forschung zu leisten.

Hinzu kommt die Notwendigkeit, Benutzerfaktoren in die TREC-Studie einzubringen. Da sich der interaktive TREC-Teil im wesentlichen mit den Auswirkungen von Relevance-Feedback-Komponenten befaßt, sind Fragestellungen wie Visualisierung komplexer IR-Funktionalität, Einbringung von softwareergonomischen Aspekten etc. derzeit nicht inhaltlich vertreten. Auch die Erarbeitung der Anfragen im Test wird derzeit – wie oben erwähnt – durch die Entwickler vorgenommen, die über bestes Wissen im Hinblick auf die Funktionalität verfügen. Dies hat mit realen Situationen im späteren Einsatz nur wenig gemein.

Ein weiteres Problem ist in der Lückenhaftigkeit der TREC-Studie zu sehen, d.h. daß keine Vollständigkeit bzgl. der potentiellen IR-Funktionalität bzw. möglicher kombinatorischer Ansätze vorhanden ist. Die Teilnehmer entscheiden in jeder Phase darüber, mit welcher Systemvariante sie teilnehmen. Aus diesem Grund ist es schwierig, generelle Fragestellungen wie z.B. „Lohnt sich die Kombination von NLP-Techniken mit statistischen Verfahren?“ zu beantworten.

4 Fazit

Derzeit bestehen zwischen den dargestellten methodischen Ansätzen noch sehr wenig Anknüpfungspunkte, um unmittelbare Verbindungen aufzubauen. Jedoch ist die TREC-Methode kein starres Konzept, sondern flexible genug, um relevante, neue Fragestellungen einbringen zu können. Durch die Methode der Erfolgsfaktorenanalyse lassen sich für verschiedene Bereiche entscheidende Kriterien erarbeiten. Die dargestellte Studie an der Universität Konstanz zeigte, daß gerade Faktoren besondere Bedeutung zukommt, die derzeit bei TREC noch vernachlässigt werden. Längerfristiges Ziel muß es sein, daß sich Studien wie TREC an diesen Faktoren orientieren, um ihre Aussagen den tatsächlichen Bedürfnissen anzupassen. Ein erster Schritt wäre in der Einbeziehung realer Benutzer zu sehen, um auch in diesem Punkt Realitätsnähe zu beweisen. Auch in bezug auf das Spektrum der TREC-Nebenstränge würde sich eine Integration der Bewertung von Benutzungsschnittstellen, Visualisierungskomponenten und multimedialen Interaktions- und Dialogformen anbieten. Dies macht auch im Bereich der Bewertungsmethoden neue Wege erforderlich.

Literatur

- BAYRAKTAR, O., WOMSER-HACKER, CH. (1998), Erfolgsfaktorenanalyse von Information-Retrieval-Systemen. Arbeitsbericht. Lehrstuhl Informationsmanagement, Universität Konstanz.
- BOLLMANN, P. (1977), Untersuchung von Effektivitätsmaßen für Dokumenten-Retrieval-Systeme. Berlin.
- BOLLMANN, P. (1981), Metrische Verfahren des Vergleichs und der Auswahl von Datenbanken. In: DEUTSCHE GESELLSCHAFT FÜR DOKUMENTATION E.V. (ED.), Deutscher Dokumentartag 1980. München et al.
- BOLLMANN, P., CHERNIAVSKY V.S. (1981), Measurement-theoretical Investigation of the MZ-metrics. In: ODDY, R.N., ROBERTSON, S.E., VAN RIJSBERGEN, C.J., WILLIAMS, P.W. (EDS.), Information Retrieval Research, London et al.
- HARMAN, D. (1998), The Text Retrieval Conferences (TREC) and the Cross-Language Track. In: RUBIO, A., GALLARDO, N., CASTRO, R., TEJADA, A. (EDS.), Proceedings of the First International Conference on Language Resources & Evaluation, Granada, pp. 517-522.
- HARMAN, D. (1995), The TREC Conferences. In: KUHLEN, R., RITTBERGER, M. (EDS.), Hypertext – Information Retrieval – Multimedia. Synergieeffekte elektronischer Informationssysteme. Proceedings HIM 95, pp. 5-28.
- HEINRICH, J.-L. (1996), Informationsmanagement. Planung, Überwachung und Steuerung der Informationsinfrastruktur. München.
- HORVÁTH UND PARTNER (EDS.) (1997), Qualitätscontrolling. Ein Leitfaden zur betrieblichen Navigation auf dem Weg zum TQM. Stuttgart.
- KRAUSE, J., WOMSER-HACKER, CH. (HRSG.) (1990), Das Deutsche Patentinformationssystem. Entwicklungstendenzen, Retrievaltests und Bewertungen. Köln et al.
- SALTON, G. (1971), The SMART Retrieval System – Experiments in Automatic Document Processing. Englewood Cliffs, NJ.
- VAN RIJSBERGEN, C.J. (²1979), Information Retrieval. London et al.
- VOORHEES, E., HARMAN, D. (1998), Overview of the sixth Text Retrieval Conference (TREC-6). Proceedings of the Sixth Text Retrieval Conference (TREC-6). In press.

Ein Navigationsassistent für das World Wide Web

Bernard Bekavac, Marc Rittberger

Informationswissenschaft,
Universität Konstanz
D-78457 Konstanz
{Bernard.Bekavac, Marc.Rittberger}@uni-konstanz.de

Inhalt

Einleitung

Graphische Übersichten

Dialoghistorien

Andere relevante Arbeiten

Der Navigationsassistent

Literatur

Zusammenfassung

Die Orientierung der Anwender während der Erarbeitung von Informationen im WWW wird auch von den neuesten Versionen aktueller Web-Browser nur unzureichend unterstützt. Als Lösungsansatz wird in dieser Arbeit eine graphische interaktive Visualisierung der Navigationsstrukturen während der Navigation vorgeschlagen. Nach einer Einführung in die Problematik der graphischen Übersichten im Hypertextbereich wird auf die Notwendigkeit neuer Arten von Dialoghistorien hingewiesen. Am besten geeignet für die Darstellung von Dialoghistorien in Netzwerken, die den Charakter von Hypertextsystemen haben, erweisen sich hierarchische Strukturen. Es wird der auf hierarchischen Strukturen basierende, im Rahmen des DeMeS-Projekt Medienassistenten entwickelte Navigationsassistent für das World Wide Web vorgestellt.

Einleitung

Die primäre Nutzung des World Wide Web in Form der Erarbeitung von Informationen durch navigatorisches Browsen in WWW-Dokumenten und die (kontextabhängige) Speicherung bzw. Administration der vom Nutzer als relevant befundenen Seiten werden von den aktuellen Browsern nur unzureichend unterstützt. Vor allem sogenannte Serendipity-Effekte¹ [Kuhlen 1991, S.126 ff.] innerhalb des riesigen Angebotes im WWW führen zu einem schnellen Orientierungsverlust des Anwenders. Dem wird im World Wide Web mit umfangreichen Katalogen, welche die Informationen klassifizieren und ordnen, und riesige Datenmengen bearbeitenden Suchmaschinen begegnet. Auf seiten der WWW-Browser wird höchstens Unterstützung in Form von Bookmarks und temporären Listen besuchter Seiten gegeben.

Im DeMeS-Projekt Medienassistenten² wurde ein Lösungsansatz zu den genannten Problemen erarbeitet. Über eine graphische interaktive Visualisierung der Navigationsstrukturen während der Erarbeitung bestimmter Informationen wird dem Anwender zur jeder Zeit eine globale Übersicht seiner Aktivitäten angeboten und dadurch eine wesentlich bessere Orientierung ermöglicht. Die technische Realisierung wurde mit dem Ziel der Plattformunabhängigkeit, vor allem aber auch Browserunabhängigkeit, verfolgt.

Graphische Übersichten

Graphische Übersichten dienen dazu, einen einfachen Überblick über ein Hypertextnetzwerk zu geben. Sie werden genutzt, um Abkürzungen zwischen weit auseinanderliegenden Knoten zu finden oder Orientierungsverluste zu beseitigen. Sie sind kognitiv leicht zu behandeln und dienen dem Leser dazu, Fixpunkte zu identifizieren und daraus eine Pfadkarte zu erstellen [Hammwöhner 1997, S.228]. Unterscheiden kann man zwischen inhaltlichen Übersichten, die sich auf die durch Knoten und Verknüpfungen basierende Struktur eines Hypertextes beziehen, und zeitlichen Übersichten, wie die Dialoghistorie, welche den Navigationsverlauf eines Benutzers im Hypertext beschreiben.

Eine gängige Form, Strukturen in graphischen Übersichten zu visualisieren, sind Hierarchien. Hierarchien eignen sich in einer graphischen Übersicht besonders, um Komplexität zu reduzieren, da sie grundlegenden mentalen und kognitiven Prinzipien entsprechen, wie Menschen Objekte organisieren, und somit auch das Navigationsverhalten unterstützen ([Parunak 1989], [Roppel 1998]). Eine Hierarchie oder Baumstruktur entspricht auch den Verhaltensweisen bei der Navigation selbst, welche Verzweigungspunkte enthält, von denen aus immer neue Navigationspfade eingeschlagen werden.

¹ Serendipity bezeichnet die Ablenkung durch ein interessantes Thema, auf das der Anwender zufällig während der gezielten Suche nach anderen bestimmten Informationen gestoßen ist. Dieser Mitnahmeeffekt kann so intensiv sein, daß das ursprüngliche Ziel ganz aus dem Auge verloren wird.

² <http://www.medienassistent.org>

[Utting & Yankelovich 1989] beschreiben die Probleme von graphischen Übersichten, welche überwiegend mit der Darstellung einer hohen Anzahl von Verknüpfungen und Knoten verbunden sind. Graphische Übersichten, die wie bspw. Dialoghistorien ständig anwachsen, erreichen solche kritischen Größen, so daß für ihr Design besondere Kriterien angesetzt werden müssen ([Mukherjee & Foley 1995], [Bekavac & Rittberger 1997]):

- Das Layout des Baumes muß kognitiv einfach sein
- Der Knoteninhalt muß für den Benutzer erkennbar sein
- Der hierarchische Baum muß den zur Verfügung stehenden Ausgabebereich optimal nutzen, weder sollten zu viele Objekte enthalten sein, noch sollten die Objekte eine bestimmte zur Identifikation notwendige Größe unterschreiten
- Der Benutzer muß in der Lage sein, den Baum zu überblicken und sich in ihm zu bewegen

Dialoghistorien

Um zeitliche Übersichten oder Dialoghistorien optimal auf ihren Einsatz vorzubereiten, ist es wichtig, über das Navigationsverhalten von Benutzern in Hypertexten Erfahrungen zu sammeln. Im World Wide Web haben verschiedene Untersuchungen gezeigt, daß sich die Benutzer bei der Navigation in der Dialoghistorie vorwiegend mittels des Back-Buttons bewegen, welcher die zuletzt vorgenommenen Navigationsschritte zurücknimmt. [Tauscher & Greenberg 1997] untersuchten über sechs Wochen das Navigationsverhalten von 23 Benutzern. Dabei zeigte sich, daß

- 58% der besuchten Webseiten erneute Besuche dieser Seite sind,
- 43% wurden bei den letzten 10 Navigationsschritten schon einmal angesteuert
- bei 15% der Seiten wurden seit dem letzten Besuch mehr als 10 Navigationsschritte getätigt
- es wurden eher die zuletzt besuchten Seiten nochmals aufgesucht
- es wurden eher nur wenige, ausgesuchte Seiten häufiger besucht
- es gibt Verzweigungsseiten, die als Ausgangspunkte regelmäßig ausgewählt wurden
- die Navigation bewegte sich eher in die Tiefe, denn in die Breite
- es wurden ständig neue Seiten besucht, die in das Repertoire der Benutzer aufgenommen wurden

Bei einer ähnlichen Studie [Catledge & Pitkow 1995] mit 107 Benutzern und drei Wochen Beobachtungszeit zeigte es sich, daß Backbuttons bei 41% aller Navigationsschritte genutzt werden. Bei [Tauscher & Greenberg 1997] wurde der Backbutton bei 30% aller Navigationsschritte eingesetzt. Während Bookmarks, neben Suchmaschinen und Katalogen, verstärkt bei der Erkundung

oder dem Einstieg ins World Wide Web benutzt werden [GVU 1996], finden Dialoghistorien kaum Anwendung. Obwohl ein Rücksprung über mehrere Seiten mit dem Back-Button sehr unkomfortabel ist, wird die dafür vorgesehene Dialoghistorie gängiger WWW-Browser nur wenig benutzt. Gründe dafür können darin liegen, daß für dieses Werkzeug prinzipiell keine Verwendung bzw. der Aufbaumechanismus der eindimensionalen Dialoghistorie nicht ersichtlich ist, oder daß bisherige Ausführungen des Werkzeuges nicht dem erwarteten und für die Benutzung notwendigen Leistungsstand entsprechen. Die neu gestaltete Visualisierung besuchter Web-Seiten der aktuellen Internet Explorer Version (ab der Version 4.0) ist jetzt zwar grafisch ansprechend aufbereitet, der zeitliche Dialogverlauf jedoch ist nun gar nicht mehr zu erkennen. Abgebildet werden dort besuchte WWW-Seiten gruppiert nach Servern und innerhalb dieser alphabetisch sortiert nach den Dokumenttiteln.

Als Lösungsansatz zu dem genannten Problem wurde im Rahmen des DeMeS-Projektes „Medienassistenten“ eine eigene Dialoghistorie für das WWW konzipiert. Dabei konnte auf Erfahrungen und Überlegungen bei der Erstellung einer Dialoghistorie für das Konstanzer Hypertext System [Hammwöhner 1997, S.184] zurückgegriffen werden. Mit dem vorhandenen Wissen und unter Beachtung bisheriger Arbeiten im Kontext des WWW zum Aufbau von Dialoghistorien ([Dömel 1994], [Ayers & Stasko 1995], [Hirsch et al. 1997], s. auch 4.) wurden folgende Randbedingungen festgelegt:

- Hierarchische Darstellungen sind für Benutzer leicht verständlich und erlauben ihnen, die wesentlichen Zusammenhänge problemlos zu erkennen
- Eine dem Benutzer vertraute graphische Darstellung sollte verwendet werden, damit er nicht mit Neuheiten belastet wird, welche ihm keinen Mehrwert bieten. Eine vertraute Darstellung erlaubt dem Benutzer, sofort mit dem Werkzeug zu arbeiten, ohne daß er bspw. mit Orientierungsverlusten rechnen muß
- Übersichtlichkeit wird bspw. erreicht durch eine vertraute graphische Darstellung. Sie ist natürlich aber auch dadurch erreichbar, daß nicht zuviel Information auf dem Bildschirm präsentiert wird, sondern nur die im gegebenen Kontext für den Benutzer relevante Information unter softwareergonomischen Rahmenbedingungen präsentiert wird
- Das Ein- und Ausblenden von Teilbäumen erlaubt es, die Komplexität der Darstellung zu verringern und dadurch die Übersichtlichkeit zu wahren. Sie ist ein vertrautes Mittel und für den Benutzer leicht verständlich. Andere Reduktionsmechanismen, wie bspw. die aus der Hypertextforschung bekannten lokal/global Maps, sind in bezug auf die Rezeptionsfähigkeit des Benutzers nicht so gut geeignet wie das Ein- und Ausblenden von Teilbäumen
- Die kontextabhängige Darstellung der einzelnen graphischen Symbole ist notwendig, um den aktuellen Zustand bspw. der Ein- und Ausblendungen zu zeigen. Die kontextabhängige Darstellung einzelner Knoten kann den Benutzer erst den wahren Umfang des betrachteten Baumes erahnen lassen, bspw. durch eine verschiedenartige Färbung, je nachdem wieviele

Knoten sich hinter der Ausblendung verbergen

- Direkte Interaktion ist unumgänglich, um eine unmittelbare Interaktion des Benutzers mit dem System zu erreichen. Der Benutzer muß im Baum navigieren können und über den Baum einzelne Seiten des Hypertextes, dessen Übersicht visualisiert wird, unmittelbar anwählen können. Nur auf diesem Wege läßt sich ein Navigationsassistent effektiv bedienen, so daß er eine hohe Benutzerakzeptanz erfährt.
- Zusatzinformationen sind notwendig, um den Benutzer Informationen über die einzelnen Knoten zu geben, so daß er eine Vorstellung erhält, was ihn erwartet, sofern er zu dem Knoten hin navigiert. Das können Hinweise auf den Titel bzw. Autoren, eine Zusammenfassung oder auch formale Hinweise, bspw. wann die Seite das letzte Mal geändert wurde oder die Adresse (URL) der Zieleinheit sein
- Plattform- und Browserunabhängigkeit ist eine technische Notwendigkeit, die als Anforderung für Software, welche über das Internet benutzbar sein soll, unabdingbar ist.

Andere relevante Arbeiten

Webmap [Dömel, 1994] war eines der ersten Programme, welches Dialoghistorien im WWW grafisch visualisierte. Dabei werden zur Laufzeit Informationen über die Navigationspfade des Anwenders aufgenommen und daraus ein zweidimensionaler Navigationsbaum gebildet. Das in [incr TCL] geschriebene eigenständige Programm (Abb. 4.1) arbeitete nur mit den inzwischen kaum mehr verwendeten Browsern Mosaic und tkWWW zusammen. In der graphischen Repräsentation von Webmap werden die Knoten, also die Web-Seiten, entsprechend der Zugriffsreihenfolge durchnummeriert. Wurde eine Seite über mehrere Wege annavigiert, so wird die Seite dem ersten Ausgangsknoten zugeordnet, die weiteren Verknüpfungen werden über gestrichelte Linien symbolisiert. Beschriftet werden die Knoten mit ihren Nummern, in der Statuszeile werden die zugehörige URL und der Dokumenttitel angezeigt. Problematisch wird die gewählte Darstellungsform bei komplexeren Bäumen. Zum einen können die überkreuzenden Kanten dann nur schwer entflochten werden, zum anderen sind die Knoteninformationen erst auf den zweiten Blick ersichtlich, da die Knoten mit ihrer Nummer und nicht dem tatsächlichen Dokumenttitel beschriftet sind.

Ein weitere Visualisierung der Dialoghistorie während der Navigation im WWW ist im WWW-Browser des Konstanzer Hypertext Systems (KHS) integriert [Bekavac, 1995]. Der Browser beinhaltet zur Navigationsunterstützung des Anwenders einen graphischen, interaktiven und mit diversen Funktionen versehenen History-Browser (Abb. 4.2). Beschriftete Knoten (Rechtecke), die bei der Knotenentstehung automatisch die Bezeichnung eines WWW-Dokuments beinhalten, stellen besuchte WWW-Dokumente dar. Der Aufbau der Knotenhierarchie folgt exakt den Navigationsschritten, bei jeder neu erreichten Web-Seite wird ein Knoten erstellt und eingefügt. Das aktuelle WWW-Dokument, das zu dem Zeitpunkt auf der KHS-Benutzerschnittstelle zu sehen ist, wird markiert dargestellt. Die Selektion eines Knotens im History-Browser

bewirkt den Wechsel vom aktuellen WWW-Dokument zu dem selektierten Dokument. Wählt man beim Navigieren im Web-Browser ein Dokument an, das schon einmal innerhalb des aktuellen Kontextes (Baumes) besucht wurde, so entfällt eine erneute Übertragung und es wird innerhalb des History-Browsers direkt zu dieser WWW-Seite verzweigt. Wird ein neuer Ausgangspunkt im WWW gewählt, so wird im History-Browser eine neue hierarchische Struktur eingefügt, die sich durch entsprechende Navigation innerhalb ihrer WWW-Dokumente entfaltet. Zusätzlich läßt sich bei dem History-Browser die Dialoghistorie drucken, die Anordnung der hierarchischen Strukturen verändern, der Knotentext editieren und die Baumstruktur so verändern, daß der Anwender nach seinen eigenen Kriterien WWW-Dokumente bzw. Gruppen von WWW-Dokumenten zusammenstellen kann. Der 1994 entwickelte Web-Browser von KHS ist bedingt durch die Weiterentwicklung von HTML in Richtung Tabellen, Frames, animierten Bildern, Java usw. nicht mehr anwendbar. Das Konzept und die Funktionalität des History-Browsers hinsichtlich der Unterstützung des Benutzers während der Navigation findet sich jedoch in vielen Bereichen des in dieser Arbeit entstandenen Navigationsassistenten wieder.

Auch MosaicG [Ayers/Stasko, 1995] ist eine browserabhängige Visualisierung der WWW-Dialoghistorie (Abb. 4.3). Die Implementierung dieses Navigationstools erfolgte direkt in den Source-Code von NCSA Mosaic in der Version 2.5. Damit ist die Nutzung auch nur auf diese Version des Browsers beschränkt. Das Grundprinzip ist gleich dem von Webmap. Werden im Browser neue Dokumente annavigiert, bindet MosaicG diese automatisch in die Baumstruktur ein. Die Autoren bezeichnen dieses Vorgehen als „reflective approach“. Zu jedem Knoten wird der Titel, die URL sowie eine verkleinerte Screen-Shot (Thumbnail) der zugehörigen Web-Seite dargestellt.

Im Gegensatz zu den bisher aufgeführten Arbeiten wurde bei dem HistoryGraph [Hirsch et al. 1997] schon eher an Browserunabhängigkeit gedacht. Implementiert wurde HistoryGraph (Abb. 4.4) in Tcl/Tk. Das Tool arbeitet zusammen mit dem Netscape Navigator unter Windows NT und mit NCSA Mosaic unter verschiedenen UNIX-Systemen. Die Verbindung zu den Browsern basiert unter Windows NT auf DDE (Dynamic Data Exchange) und unter UNIX auf CCI (Common Client Interface). Jeder Knoten ist mit dem Dokumenttitel bzw. der URL der zugrundeliegenden Webseite beschriftet. Um Platz zu sparen, wird diese Beschriftung abgekürzt. Über ein vorangestelltes Icon werden zusätzliche Informationen visualisiert. In Statuszeilen werden zum selektierten Knoten der Dokumenttitel, die URL sowie die Anzahl der Besuche angezeigt. Der Aufbau des History-Graphen entspricht in etwa dem beim KHS-History-Browser.

Das Programm HistoryTree ist ein Shareware-Programm der Firma Smartbrowser³. Es unterstützt die Browser Netscape Navigator/Communicator (ab der Version 2) sowie den Microsoft Internet Explorer (ab der Version 3). Auch hier ist der Aufbaumechanismus ähnlich den anderen vorgestellten

3 Die Shareware-Version der Software ist unter <http://www.smartbrowser.com> verfügbar

Lösungen. Die Art der Visualisierung ist im Vergleich zu den anderen etwas unübersichtlich (Abb. 4.5). Die Knoten sind nicht beschriftet, wodurch eine schnelle Orientierung im Baum unmöglich ist. Manuell kann zu jedem Knoten eine Bezeichnung eingegeben werden, welche dann im Baum rechts vom Knoten angezeigt wird. Informationen zu einem Knoten werden als QuickInfo angezeigt, wenn der Mauszeiger über dem Knoten positioniert wird. Zusätzlich ist es möglich, zu einem Knoten einen kleinen Screenshot in Icongröße anzufertigen, dessen Nutzen wegen der Größe und Anordnung (Überlappung) etwas zweifelhaft ist.

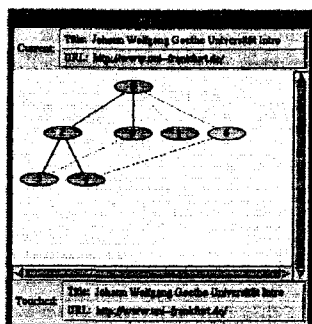


Abb. 4.1: Webmap

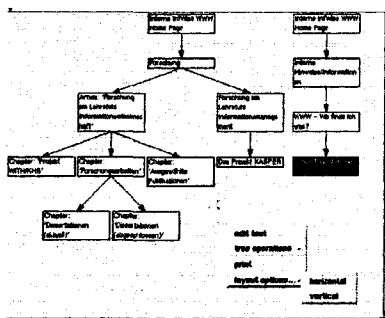


Abb. 4.2: KHS-WWW-History-Browser

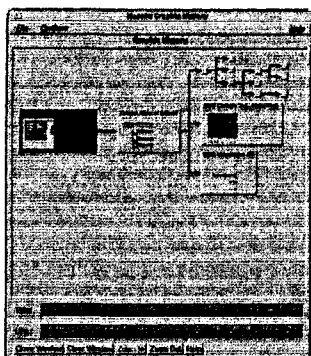


Abb. 4.3: MosaicG

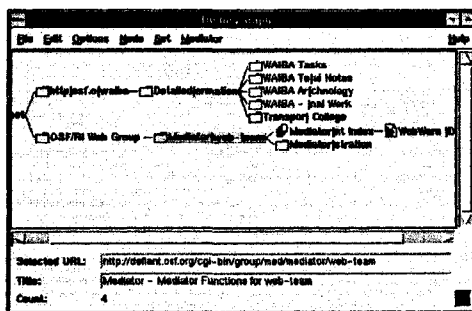


Abb. 4.4: HistoryGraph

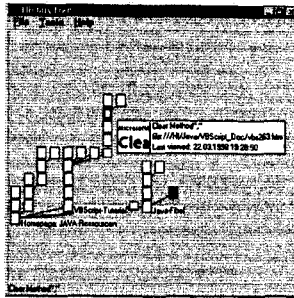


Abb. 4.5: HistoryTree

Der Navigationsassistent

Als eine erste softwaretechnische Realisierung des in dieser Arbeit entworfenen Konzepts einer Dialoghistorie für das WWW wurde der Navigationsassistent entwickelt (Abb. 5.1). Bei diesem wird über eine graphische, interaktive Visualisierung der Strukturen während des Navigierens dem Anwender zur jeder Zeit eine globale Übersicht seiner Aktivitäten angeboten und dadurch eine bessere Orientierung ermöglicht. Um dabei die Plattform- bzw. Browserunabhängigkeit zu gewährleisten, wurde die Visualisierung der Dialoghistorie selbst, als HTML-Seite in Verbindung mit der Scriptsprache *JavaScript*, die von allen gängigen WWW-Browsern unterstützt wird, realisiert.

Der Kernpunkt der Implementierung war die Unterstützung der Orientierung des Benutzers durch geeignete Visualisierung der Navigationspfade während der Navigation im WWW. Je nach Interesse an ernavigierten Dokumenten bzw. der Relevanz dieser, bewegen sich Anwender, ausgehend von einer Startseite, von Dokument zu Dokument und bilden dabei einen individuellen Navigationspfad. Bewußt, z.B. über den Back-/Forward-Button des Browsers, aber auch oftmals unbewußt, z.B. über zyklische Verknüpfungen zwischen WWW-Dokumenten, wird dabei innerhalb dieses Pfades immer wieder vor- und zurücknavigiert, oder es werden von schon besuchten Dokumenten aus neue Pfade eingeschlagen. Solche „Verästelungen“ lassen sich gut mit Hilfe gerichteter Graphen darstellen. Definiert man dabei die Rückkehr zu schon besuchten Dokumenten nicht als eine Kante zu einem neuen oder schon vorhandenen Knoten, sondern als einen einfachen Focus-Wechsel zu dem entsprechenden Knoten innerhalb des Graphen, so wird der Graph azyklisch, dadurch hierarchisch, und in Form einer Baumstruktur sinnvoll darstellbar. Diese Analogie entspricht auch im wesentlichen dem Aufbau von WWW-Hypertexten, deren hierarchische Strukturen zwar nicht explizit formuliert, dennoch ausgehend von einer Startseite und unter Betrachtung eines bestimmten Kontextes vorhanden sind. Somit werden, beginnend mit der Wurzel des Baumes, die das erste besuchte WWW-Dokument darstellt, alle anderen besuchten Dokumente als Knoten bzw. Blätter des Baumes dargestellt.

Es wurden verschiedene Ansätze zur Darstellung und Anordnungsmöglichkeiten solcher Baumstrukturen auf Relevanz für das genannte Problem

geprüft und diskutiert. Neben klassischen Darstellungen wurden z.B. auch Cone-Trees oder Hyperbolic-Trees in Erwägung gezogen:

- In Cone-Trees werden die Informationseinheiten in Kegelform unter dem Ausgangsknoten dargestellt. In [Hemmje 1993] werden umfangreiche Hierarchien durch miteinander verbundene, rotierende Zylinder dargestellt.
- Hyperbolic-Trees dienen dazu, Hierarchien mithilfe des Fish-Eye-Prinzips darzustellen. Die Knoten im Zentrum sind vollständig dargestellt, während zum Rand hin die Darstellung immer weniger detailliert ist und einzelne Knoten und Kanten sich überlappen.

Eine Untersuchung mit 14 Teilnehmern zeigte, daß die Hyperbolic Trees beim Lösen von Suchproblemen mittels Navigation in Hierarchien keine Vorteile in der Performanz gegenüber klassischen hierarchischen Darstellungen erbrachten. Die Teilnehmer bevorzugten eindeutig die ihnen zum Vergleich zur Verfügung gestellte klassische Darstellung, wobei die Hyperbolic Trees bei den Kriterien neue Technologien, interessante Grafik oder Einzigartigkeit leichte Vorteile hatten [Czerwinski & Larson 1997].

Eine naheliegende Lösung für die Darstellung der klassischen Anordnung wäre die Verwendung der ebenfalls im Projekt Medienassistenten mitentwickelten Klassen zur Visualisierung von Suchergebnissen mittels gerichteten azyklischen Graphen [Bekavac & Rittberger 1997]. Diese Klassen wurden in Java geschrieben und würden somit ebenfalls eine Plattform- bzw. Browserunabhängigkeit sichern. Da aber im vorliegenden Fall des Navigationsassistenten der Anwender, im Gegensatz zu der Suchergebnisvisualisierung, jederzeit mit der entsprechenden Grafik arbeiten und interagieren muß, wurde nach einer "bekannten" Darstellungsart gesucht, die vom Anwender intuitiver verwendet und somit schnell akzeptiert werden würde.

Im Stil graphischer Repräsentation von Dateiverzeichnissen gängiger Benutzeroberflächen verschiedener Plattformen wie Windows, Unix oder Macintosh wurde solch eine Darstellungsart gefunden. Gegenüber einer anderen Baumdarstellung ist der größte Vorteil das „look and feel“ dieser Visualisierung, das dem Großteil der Internet-Nutzer bekannt ist. Weiterhin können bei dieser Anordnung durch direkte Manipulation Teilbäume ein- und ausgeblendet werden, wodurch die Komplexität des angezeigten Baumes benutzerspezifisch verringert werden kann und der Baum auch bei einer größeren Knotenmenge übersichtlich bleibt.

Ein weiterer Vorteil dieser weit verbreiteten Metapher waren im Internet schon vorhandene Software-Bausteine, mit denen sich Anordnungen dieser Art einfacher und schneller realisieren lassen. Dabei konnte auf die bereits für das Internet entwickelte Basis-Software „Joust“ der Firma *Alchemie Computing Ltd.* zurückgegriffen werden, die im WWW frei zur Verfügung steht. Die in JavaScript geschriebene Software eignet sich zur Darstellung von Übersichten von Web-Sites, sogenannten „Site-Maps“, im bereits erwähnten Stil. Dabei paßt sich die detaillierte Darstellung (Icons usw.) automatisch der Plattform des verwendeten Web-Browsers an und erfüllt somit die Forderung nach einer dem

Benutzer vertrauten Oberfläche. Diese Software wurde auf die für den Lösungsansatz benötigte dynamische Darstellung von Navigationsstrukturen umgeschrieben bzw. erweitert.

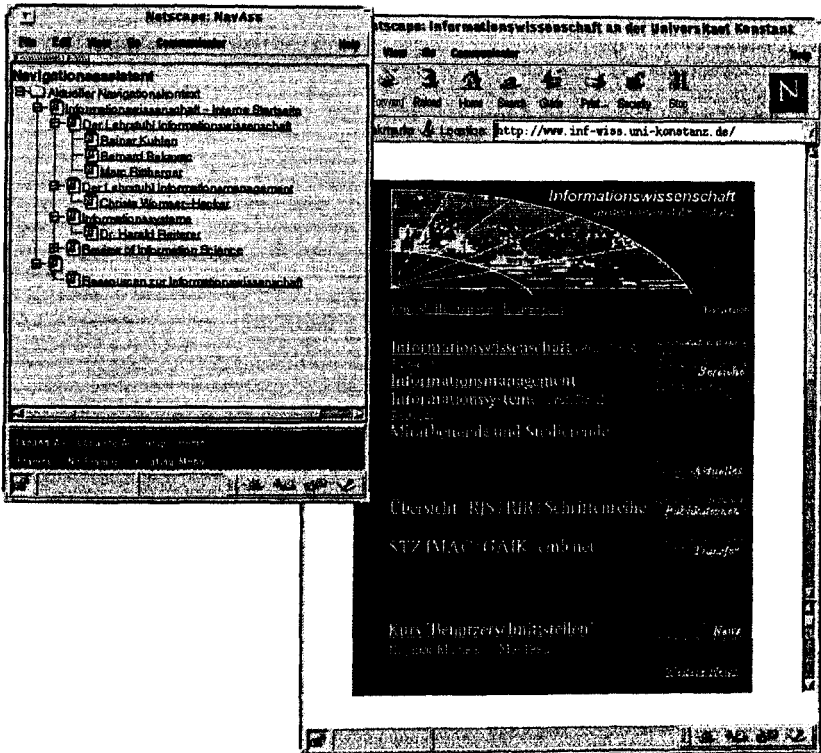





Abb. 5.1: WWW-Browser und Navigationsassistent mit der Darstellung der aktuellen Dialoghistorie






Um den Navigationspfad des Anwenders möglichst originalgetreu nachzubilden, wurde folgende Logik beim Aufbau des Navigationsbaumes zugrundegelegt:

- Baum-Knoten in textueller Form stellen besuchte WWW-Dokumente dar
- Das aktuelle WWW-Dokument, das zu dem Zeitpunkt auf dem Browser zu sehen ist, wird gesondert markiert dargestellt (im Navigationsassistenten der Abb. 5.1 als heller und vergrößerter Text dargestellt)
- Die Selektion eines Eintrags (Knotens) innerhalb des Navigationsbaumes bewirkt den Wechsel vom aktuellen WWW-Dokument zu dem neu selektierten Dokument im Browser-Fenster

- Immer dann, wenn innerhalb des aktuellen WWW-Dokuments im Browser über einen Hypertextlink zu einem neuen WWW-Dokument navigiert wird, wird unterhalb des zugehörigen Eintrags im Navigationsbaum ein neuer Eintrag, der als Beschriftung den Titel des Dokuments trägt, angefügt und markiert dargestellt
- Wählt man beim Navigieren ein Dokument an, das schon einmal innerhalb des aktuellen Kontextes (Baumes) besucht wurde, so wird kein neuer Knoten eingefügt, sondern es wird innerhalb des Navigationsbaumes der entsprechend vorhandene Eintrag markiert. Dies gilt auch für Dokumente, die durch Benutzung der Navigationshilfen des Browsers, wie Back/Forward oder der History-Liste, erneut geladen werden
- Wird ein neuer Ausgangspunkt im WWW gewählt, z.B. durch die direkte Eingabe einer URL in das Adreßfeld des Browsers, so wird im Navigationsassistenten eine neue hierarchische Struktur (neuer Navigationskontext) eingefügt, die sich durch entsprechende Navigation innerhalb ihrer WWW-Dokumente entfaltet

Um den Benutzer bei seiner Navigation bestmöglich zu unterstützen, werden ihm neben dem Titel der besuchten Web-Seite weitere Informationen zu diesem Dokument zur Verfügung gestellt. Jeder Knotenbeschriftung wird ein Sinnbild (Icon) vorangestellt, um den Benutzer darüber zu informieren, was sich hinter diesem Knoten verbirgt. Entsprechend wurde bei der Konzeption des Navigationsassistenten dieser informationelle Mehrwert berücksichtigt. Folgende Symbole wurden verwendet, um den Benutzer mit zusätzlichen Informationen über das besuchte Hypertext-Dokument zu versorgen:

 Dieses Icon symbolisiert einen Knoten und somit ein WWW-Dokument. Je nachdem, welcher Browser und welche Plattform verwendet wird, kann die Darstellung variieren: Bei Verwendung des Netscape Communicators wird immer das Symbol  angezeigt, unter OS/2 wird bei anderen Browsern das Icon  verwendet.





  Mit diesen Icons werden geschlossene und offene Teilbäume symbolisiert. Die dem Knoten zugehörigen Komponenten, sofern vorhanden, werden entsprechend angezeigt. Gleichzeitig wird durch das Icon die Informationen vermittelt, daß von dem Hypertext-Dokument, welches durch den Knoten repräsentiert wird, weitere Navigationsschritte zu anderen Hypertext-Dokumenten erfolgt sind⁴. Bei Verwendung unterschiedlicher Plattformen werden automatisch plattformspezifische Symbole verwendet: Neben den schon abgebildeten Icons für Unix und Window, gibt es noch unterschiedliche Darstellungen unter Macintosh () und OS/2 (, ).

Ein weiteres Konzept, welches für den Überblick in umfangreichen Bäumen wichtig ist, ist die Möglichkeit, durch Interaktion mit der Maus die Blätter oder Knoten von Teilbäumen aus- bzw. wieder einblenden zu können. Für diese

⁴ Aus softwaretechnischen Gründen wird dieses Symbol momentan nur für die Unterscheidung verschiedener Navigationskontexte verwendet.

Funktion ist vor dem Icon eines Knotens evtl. eine Checkbox vorhanden. Im Falle solch einer Checkbox existieren weitere Knoten oder Blätter als Komponenten des repräsentierten Knotens. Fehlt die Checkbox, existieren keine weiteren Komponenten.

Die Symbole für die Checkboxes haben folgende Darstellung:

- ☒ Es existieren weitere Komponenten zu diesem Knoten, die aber momentan nicht angezeigt werden. Mit einem einfachen Klick der linken Maustaste auf das Checkboxsymbol ☒ werden die zugehörigen Komponenten angezeigt und das Checkboxsymbol wechselt zu ☐. Das für die Repräsentation des Knotens verwendete Icon wechselt im Fall des Navigationsassistenten von  auf .
- ☐ Die Komponenten des Knotens sind angezeigt. Durch einen einfachen Klick der linken Maustaste auf das Checkboxsymbol ☐ werden die dem Knoten zugehörigen Komponenten wieder verborgen. Das Checkboxsymbol wechselt zu ☒. Das für die Repräsentation des Knotens verwendete Icon wechselt dann von  auf .

Exakt dieselbe Funktionalität kann durch einen Doppelklick der linken Maustaste auf das Icon erreicht werden. Je nachdem, ob der Teilbaum gerade geöffnet ist oder nicht, bedeutet der Doppelklick das Aus- oder das Einblenden der Komponenten dieses Teilbaumes. In der Verwaltung der graphischen Darstellung wird der Status (ein- bzw. ausgeblendet) aller Komponenten gespeichert. So können die Zustände der darunter liegenden Teilbäume (geöffnet bzw. geschlossen) wiederhergestellt werden.

Die Art der Implementierung der Software läßt in zukünftigen Versionen noch weitere Darstellungen und Metaphern im Kontext der ernavigierten Web-Seite und der Dialogstruktur zu.

Neben der grundsätzlichen Information über besuchte Web-Seiten und die Navigationspfade zu diesen Seiten, sind zusätzliche Daten zum jeweiligen HTML-Dokument vorhanden⁵:

- der Titel des WWW-Dokumentes
- die Adresse (URL) der Web-Seite
- die URL des Dokumentes im WWW, von dem aus die gerade angezeigte Seite aufgerufen wurde (Referer)
- das Datum der letzten Veränderung des Hypertext-Dokumentes
- die Anzahl der Verknüpfungen im gerade angezeigten Hypertext-Dokument

Weitere Funktionen, wie die Möglichkeit, die Beschriftungen der Einträge benutzerspezifisch zu editieren, Teilbäume wegen mangelnder Relevanz zu löschen oder im umgekehrten Fall in eigenständige Bäume (Navigationskontexte) innerhalb des Navigationsassistenten zu verwandeln,

⁵ In der aktuellen Version werden nur Titel und URL angezeigt.

und die Möglichkeit, die Navigationsstrukturen dauerhaft zu speichern, befinden sich momentan in Bearbeitung und sollen den Navigationsassistenten praxisrelevant verbessern.

Schwierig wird jedoch die Umgehung von vorhandenen Einschränkungen sein, die durch HTML und Sicherheitsrestriktionen beim Zugriff auf verschiedene Server durch *Java* und *JavaScript* gegeben sind. Bei Dokumenten, welche in Frames aufgeteilt sind, wird im Navigationsassistenten nur die Hauptseite erfaßt. Bedingt ist dies durch Einschränkungen von *JavaScript* im Zugriff auf die einzelnen Frame-Informationen von einer Frameset-Seite aus. Dieses Problem soll laut Netscape in einer der zukünftigen Versionen von *JavaScript* behoben werden.

Ebenso problematisch ist ein Serverwechsel während der Navigation. In diesem Fall verweigert *JavaScript* den Zugriff auf interne Informationen (z.B. URL) des WWW-Dokuments einer neuen Domain, da sonst die Navigationspfade eines Benutzers seitens Dritter unbemerkt überwacht werden könnten. Solch ein "unerlaubter" Zugriff wird durch eine Sicherheitsrestriktion von *JavaScript* verhindert, die einen Domainwechsel nicht ohne Weiteres, d.h. ohne die Zustimmung des Benutzers, zuläßt. Um aber einen Benutzer diesbezüglich softwaretechnisch befragen zu können, muß ein in *JavaScript* geschriebenes Programm erst durch eine offizielle Stelle zertifiziert und dieses Zertifikat von den Browsern akzeptiert werden⁶. Solche Zertifikate sind gegen eine Gebühr von verschiedenen Zertifizierungsagenturen, wie z.B. Verisign⁷, zu beziehen.

Literatur

[Ayers & Stasko 1995]

E. Z. Ayers / J. T. Stasko. *Using Graphic History in Browsing the World Wide Web*. In: Fourth International WWW Conference, 1995
<http://www.w3j.com/1/ayers.270/paper/270.html>

[Bekavac 1995]

B. Bekavac. *Das Konstanzer Hypertext System (KHS) als WWW-Client*. In: Proceedings Konstanzer Informationswissenschaftliches Kolloquium 1995, Universitätsverlag Konstanz
<http://www.inf-wiss.uni-konstanz.de/~bekavac/kik95.html>

[Bekavac & Rittberger 1997]

B. Bekavac / M. Rittberger. *Kontextsensitive Visualisierung von Suchergebnissen*. In: Hypertext - Information Retrieval - Multimedia '97: Theorien, Modelle und Implementierungen integrierter elektronischer Informationssysteme, 307-321, Universitätsverlag Konstanz: Konstanz, Nr. 30 Schriften zur Informationswissenschaft 1997

⁶ siehe <http://developer.netscape.com/library/documentation/signedobj/overview.html> und
http://www.suitable.com/Doc_CodeSigning.shtml

⁷ wird von den meisten Browsern akzeptiert, siehe <http://www.verisign.com>

[Catledge & Pitkow 1995]

L.D. Catledge / J.E. Pitkow. *Characterizing browsing strategies in the World-Wide Web*. In: Computer Networks and ISDN Systems (27)6, 1065-1073, 1995

<http://www.igd.fhg.de/www/www95/papers/80/userpatterns/UserPatterns.Paper4.formatted.html>

[Czerwinsky & Larson 1997]

M. Czerwinski / K. Larson. *An Initial Examination of Tree Navigation versus Hyperbolic Browsing during Search: The Microsoft Site Analyst*.

<http://www.research.microsoft.com/research/ui/marycz/bhci97.html>, British HCI 1997

[Dömel 1994]

P. Dömel. *WebMap: a graphical hypertext navigation tool*. Second International World-Wide Web Conference: Mosaic and the Web. Chicago, IL, USA, 17-20 Oct 1994, In: Computer Networks and ISDN Systems (28)1-2, 85-97, 1995

[GVU 1996]

GVU. *GVU's WWW User Surveys*. 6th Survey. Browsing Strategies.

http://www.cc.gatech.edu/gvu/user_surveys/survey-10-1996/graphs/user/Browsing_Strategies.html, 1996

[Hammwöhner 1997]

R. Hammwöhner. *Offene Hypertextsysteme. Das Konstanzer Hypertextsystem (KHS) im wissenschaftlichen Kontext*. Habilitationsschrift, Universitätsverlag Konstanz, Nr. 32 Schriften zur Informationswissenschaft 1997

[Hemmje 1993]

M. Hemmje. *Eine inhaltsorientierte, intuitive 3D Benutzerschnittstelle für Information Retrieval Systeme*. In: Information Retrieval '93. Von der Modellierung zur Anwendung. Universitätsverlag Konstanz: Konstanz, Nr. 12. Schriften zur Informationswissenschaft, 82-99, 1993

[Hirsch et al. 1997]

F.J. Hirsch / W.S. Meeks / C.L. Brooks. *Creating Custom Graphical Web Views Based on User Browsing History*. 1997 In: Sixth International WWW Conference, 1997

<http://www.camb.opengroup.org/RI/www/waiba/papers/www6/hg.html>

[Kuhlen 1991]

R. Kuhlen. *Hypertext, ein nicht-lineares Medium zwischen Buch und Wissensbank*. Springer: Berlin, 1991

[Mukherjee & Foley 1995]

S. Mukherjee / J.D. Foley. *Visualizing the World-Wide Web with the Navigational View Builder*. In: Computer Networks and ISDN Systems (27)6, 1075-1084, 1995

[Parunak 1989]

H. Van Dyke Parunak. *Hypermedia topologies and user navigation*. In: *Proceedings of the Hypertext '89 conference*. (Pittsburgh, PA, 5.-8. Nov.), 43-50, New York: ACM, 1989

[Roppel 1998]

S. Roppel. *Visualisierung und Adaption: Techniken zur Verbesserung der Interaktion mit hierarchisch strukturierter Information*. Universitätsverlag Konstanz, 33 Schriften zur Informationswissenschaft 1998

[Tauscher & Greenberg 1997]

L. Tauscher / S. Greenberg. *How people revisit web pages: empirical findings and implications for the design of history systems*. <http://www.hbuk.co.uk/ap/ijhcs/webusability/tauscher/tauscher.html> , In: *International Journal of Human Computer Studies* (47)1, 97-137, 1997

[Utting & Yankelovich 1989]

K. Utting / N. Yankelovich. *Context and orientation in hypermedia networks*. In: *ACM Transactions on Information Systems* (7)1, 58-84, 1989

Section 6

The Information Society

(Different Aspects)

SCHWIERIGKEITEN BEIM SEMIOTISCHEN BLICK AUF DIE INFORMATIONSGESELLSCHAFT

Frieder Nike

Universität Bremen, FB3, Postfach 330440, D-28334 Bremen
nike@informatik.uni-bremen.de

... a bitstream collapse of space and time.
Edward Barrett 1995

In dem Buch *Technopoly* spricht der amerikanische Pädagoge und Kulturkritiker Neil Postman davon, daß die Entwicklung der westlichen Gesellschaft den Zusammenhang zwischen der stofflichen Darstellung einer Information und ihrer praktischen Verwendung tendenziell auflöse. Ohne solch eine Verbindung aber bleibe "Information" für uns ohne Bedeutung; und was für uns keine Bedeutung besitze, existiere in gewissem Sinne – nach konstruktivistischer Auffassung nämlich (Glaserfeld 1997) – für uns überhaupt nicht. Mehr noch: ohne Kontext sei Information nicht nur nutzlos, sondern sogar schädlich. (Postman 1993)

Die Auffassung, daß Information *keine* objektive, quantitativ erfaßbare, von einem Ort zu einem anderen übertragbare, von einer Person an eine andere mitteilbare Sache sei, die Auffassung also, daß Information *nicht* reifiziert werden könne, hat sich im Zuge der Kritik an der Informationstheorie und der Fortschritte der Kognitionswissenschaft stark ausgebreitet, ja, vielleicht auf breiter Front durchgesetzt (Foerster et al. o.J., Varela 1993). Die auf Shannon und Weaver (1948) zurückgehende sog. Informationstheorie hat sich längst als eine Theorie technischer Übertragung von Signalen erwiesen, und war von ihren Autoren auch nicht anders gemeint gewesen¹.

¹ Nachdem Shannon und Weaver den Begriff der Nachricht (*message*) eingeführt haben, sagen sie: "Frequently the messages have meaning; that is, they refer to ... some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem." (Shannon & Weaver 1948) Erstens erhalten wir den Hinweis, daß Bedeutung nicht in der Nachricht selbst liegt, sondern im Verweis auf anderes. Zweitens aber hat es der Ingenieur mit "Bedeutung" nun gerade *nicht* zu tun. Information in der Nachrichtentheorie ist eine Quantität, die

Dennoch spielen quantitative Auffassungen von Information im alltäglichen Reden von Politikern, Pädagogen, Medieninformatikern u.ä. nach wie vor eine große Rolle. Die Behauptung verstummt nicht, ja, sie wird sogar wieder häufiger und in voller Überzeugung vorgetragen, daß es eine ständig anwachsende Informationsflut gäbe, daß das Wissen immer schneller verdoppelt werde, daß es immer unmöglicher erscheine, einen Überblick über ein Wissensgebiet zu bewahren – wenn nicht der Computer wäre. Denn allein mit der Hilfe von Computern könne überhaupt noch eine Hoffnung auf das Ordnen und Auffinden wichtiger Informationen gehegt werden.

Gleichzeitig werden wir mit Versprechungen überschüttet, daß das "Netz" das Wissen der Welt in Wohnungen und an Arbeitsplätze liefere und daß der Weg in die selbstredend lichte Informationsgesellschaft eine breite Autobahn ohne Kollisionsgefahr und Stau sei. Wir dürften und sollten nur keinesfalls zögern, sondern rasch aufspringen und mitfahren, Lust und Wonne seien groß, und wer früh dabei sei, habe allen anderen einen entscheidenden Vorsprung voraus. Lebenslang müßten, ja: dürften wir neuerdings lernen und es gelte so schrecklich viel an Information zu bewältigen, daß unser Leben mit dieser Aufgabe vollgepackt werden müsse.

Wie es so geht in Zeiten des Umbruchs, hören wir zwei Arten von Stimmen: begeisterte Visionen und düstere Trauer, wofür wir in unserem Fall z.B. Negroponte (1995) und Stoll (1995) aufrufen können, zwei populäre Autoren, die im engeren Sinne nicht wissenschaftlich schreiben. Die Namen könnten wir durch manche andere ersetzen, solche von originelleren und tieferen Denkern und Beobachtern. Doch gerade das Alltägliche in Diktion und Orientierung macht solche Texte reizvoll als Dokumente kulturellen Diskurses.

Längst ist die Umwälzung der technischen Infrastruktur² der Gesellschaft soweit fortgeschritten, daß neue Umgangsweisen und Praktiken in der Kultur aufscheinen. Manche beklagen das. Doch Kultur entwickelt sich nicht zum Besseren oder Schlechteren, sondern zum Anderen. Dieses Andere nennen einige nun "Informationsgesellschaft". Dagegen, gegen das Wort, zu räsonnieren, muß abgeschmackt wirken. Eine Fußnote möchte ich aber dennoch anbringen, in Sorge um wissenschaftliche Redlichkeit im Begriff³ – ist es doch üblich geworden, auch Forschungsprojekte u.dgl. auf die "Informationsgesellschaft" zu orientieren.

"Nothing could be more misleading than the claim that computer technology introduced the age of information" (Postman 1993, 61). Dies ist nun nicht eben Postmans originäre Erkenntnis, eher schon die Elizabeth Eisensteins (1983), auf deren bemerkenswerte Studie er sich auch stützt. Doch die Popularisierung solcher Einsicht durch einen Autor wie Postman ist wichtig, da wir eben leider auch in wissenschaftlichen Zusammenhängen anderes zu lesen bekommen. Das Zeitalter der Information, wenn es ein solches gibt, hat lange vor der

das ausdrücklich ausschließt, worauf es im umgangssprachlichen Begriff des Wortes "Information" ankommt.

² Gemeint ist der Einzug des Einzelplatzcomputers an vielen Arbeitsplätzen, die allgemeine Ausbildung einer Struktur digitaler Netze und der Einbau informationstechnischer Komponenten in tendenziell alle technischen Systeme.

³ ... und eingedenk der Hegelschen Anstrengung des Begriffs.

Computertechnik begonnen, mit der Druckerpresse nämlich, wie wir uns leicht denken können.

Daß nicht Information das synthetisierende Moment der Informationsgesellschaft ist, sondern daß diese Rolle – lassen wir uns überhaupt auf den Begriff einer solchen Gesellschaft ein – der Informationstechnik zukommt, darum geht es in diesem Beitrag. Ich werde mich auf Begriffe der Semiotik stützen, die erlauben, recht einfach zwischen Daten, Information und Wissen zu unterscheiden⁴. Mit dieser Unterscheidung will ich den Ort kennzeichnen, an dem der Computer eingesetzt werden kann.

Die Informationsgesellschaft ist eine Gesellschaft der Informationstechnik

Wer immer heute von der "Informationsgesellschaft" spricht – und das sind viele und oft Prominente, das Wort wird längst als Leerstelle benutzt, ohne über seinen Begriff Rechenschaft abzulegen⁵ –, täte das nicht, hätte sich die maschinelle Bearbeitung dessen, was wir Information nennen, während der Dauer einer Generation nicht stürmisch und erfolgreich in fast alle Bereiche gesellschaftlicher Praxis hinein ausgebreitet. Die Informationsgesellschaft wird jedoch, entgegen ihrer Benennung, nicht von Information zusammengehalten, sondern von deren maschineller Bearbeitung mittels Computer, also von der Informationstechnik.⁶ Und was das ist, das da bearbeitet wird, ob Information oder etwas ganz anderes, das gilt es genauer zu kennzeichnen.

Die Einleitung zum vierten einer Serie von Sammelbänden, die sich mit Hypermedien, Literatur, Kultur und Geistesströmungen befassen, beginnt Edward Barrett folgendermaßen:

The computer textualizes everything. The phrase "information superhighway" tricks us into believing that it's about delivery and movement, about access and possession, about abundance and automatic knowledge. ... As if environment did not precede it, as if place, this time, this view, this thought, my words, our world, language did not precede it. (Barrett & Redmond 1995, xi)

Man kann kaum deutlicher sagen, daß Textualisierung der Maschine, Kontextualisierung aber uns zukommt. Zeit und Raum schaffen wir durch unsere leibliche Bewegung, ohne die nichts Bedeutung gewinnt.

⁴ Den Unterschied von Wissen, Information und Daten spricht Postman (1993) ebenfalls mit erfreulicher Klarheit an.

⁵ Literatur zur Informationsgesellschaft explodiert. Es gibt zumindest eine spezielle Zeitschrift, *Information Society*. Auf eines der fundierteren Bücher sei hingewiesen: (Spinner 1998).

⁶ Im Kern besteht die Informationstechnik aus dem Computer und dessen Vernetzung. Keine Schrift über die Segnungen und Flüche der Informationsgesellschaft geht an der Rolle vorbei, die der vernetzte Computer darin spielt, auch wenn mehr an technischen Einrichtungen zur Informationstechnik hinzuzurechnen ist als das Netz und der PC.

An Beispielen macht Barrett im weiteren deutlich, daß unser Interesse als lebendige menschliche Wesen stets auf die Zusammenhänge, die Hintergründe, die Verbindungen mit erlebter Welt gerichtet sind, wenn wir von Information sprechen, nicht auf die pure Form einer Aussage, eines Bildes in irgendeiner Codierung, die so gnadenlos notwendig ist, wenn sich der Computer einmischt, der dann alles eindeutig, korrekturfähig, effizient macht und behandelt. "We want, not the dead object, but the living body in its connections to the world, connections that sustain it, give it meaning", fährt Barrett weiter fort (Barrett & Redmond 1995, xv). Und er wird manchen überraschen mit der Auffassung, daß die mediale Seite des Computers sich gerade *nicht* in Vielfalt zeige, sondern in der Reduktion des Nebeneinanders der vielen auf ein einziges Medium: im Strom der Bits brechen alle raumzeitlichen Unterschiede zusammen.

Sortieren und Kategorisieren, Duplizieren und Übertragen, alles in Anwendung auf dimensionslose Bitfolgen, das sind die technischen Operationen, für die Computertechnik taugt. Was haben sie mit Information zu tun, wenn Information daraus entsteht, daß Kontexte hergestellt und evoziert, nicht negiert und zurückgedrängt werden? Solche kontextfreien Operationen sind notwendig und müssen mit großer Effizienz ausgeführt werden, um die technische Basis für einen anderen Umgang mit Information zu schaffen. Das heißt aber gerade, daß Information selbst nicht Gegenstand der maschinellen Bearbeitung ist.

Die Informationsgesellschaft wird so genannt, weil in ihr und für ihre Reproduktion als Gesellschaft⁷ die maschinelle Bearbeitung von Informationen eine zentrale Rolle spielt, d.h. die Darstellung, Speicherung, algorithmische Veränderung und Übertragung von "Information"⁸. Wenn nun das, was für mich Information ist, weil ich es in meinen lebensweltlichen Zusammenhängen sehe, auf die angedeutete Weise maschinell bearbeitet wird, dann ist notwendigerweise zunächst einmal von eben diesen meinen lebensweltlichen Zusammenhängen abzusehen. Wie sollten diese denn in die Maschine kommen, wie sollten sie außerhalb meiner Subjektivität existieren als Objekte, auf die eine Maschine angesetzt werden kann, als Objekte, die zusammenhanglos und unverrückt vorhanden blieben, auch wenn ich tot umfiele?

Der Erfolg der Informationstechnik, d.h. der Maschinisierung von Kopfarbeit⁹, hat bis vor etwa zwanzig Jahren viele Forscher nicht in Zweifel gestürzt oder zu Protesten herausgefordert, wenn die Computermetapher herangezogen wurde, um bestimmte Tätigkeiten oder Verhaltensweisen, Funktionen jedenfalls¹⁰,

⁷ Eine Gesellschaft erhält sich in der Form, die sie besitzt, indem sie sich jahraus, jahrein reproduziert. Sie reproduziert sich durch gesellschaftliche Produktion. Diese wiederum verlangt Arbeit der Mitglieder.

⁸ Wir werden in Kürze sehen, warum wir bis hierher solch ein Eiertanz um die Verwendung des Wortes "Information" aufführen. Die Maschine hat es mit etwas Einfacherem, mit Daten nämlich, zu tun.

⁹ Genaueres könnte man hierzu bei (Nake 1992) nachlesen.

¹⁰ Spätestens an dieser Stelle, bei der Reduktion dessen, was wir tun, auf *Funktionen*, hätte lautstarker Protest seitens der Forschenden erwartet werden dürfen. Denn der Trick mit der funktionalen Betrachtung zieht nach sich, daß auch die Struktur der beiden Systeme Mensch und Computer als (nahezu) gleich angesehen wird.

des Menschen zu "erklären". Das Gehirn war ein Computer geworden, kein Zweifel daran wurde erlaubt.

Niemand außerhalb technokratischer Zirkel gäbe heute noch einen Pfifferling für eine solche Position¹¹. Das schöne Buch (Schacter 1996) faßt eindrucksvoll den Erkenntnisstand zusammen, was den Unterschied zwischen menschlichem Gedächtnis und Computerspeicher betrifft¹². Was wir in der Vergangenheit erlebt haben, bestimmt weitgehend, was wir aus unseren tagtäglichen Erlebnissen behalten. Das Gedächtnis ist viel mehr eine Aufzeichnung dessen, wie wir Ereignisse *erfahren*, als eine Repräsentation solcher Ereignisse *selbst* (Schacter 1996, 6).

Indem die sog. Informationsgesellschaft und die vermutlich hinter ihren popularisierten Werbeslogans stehenden gesellschaftlichen Tatsachen – Information Superhighway, Infobahn, Wissensmanagement – samt und sonders rückführbar sind auf die technische Bearbeitung dessen, was für uns Information ist, gewinnt sie ihre Synthesis¹³ eben nicht aus der Information, sondern aus der Informationstechnik.

In anderen, ja: in *allen* Gesellschaften gehen die Menschen miteinander und mit den Dingen und Prozessen so um, daß wir, rückblickend von heute aus gesagt, in ihren Handlungen jenes Verhältnis aufspüren können, das wir Information nennen. Diese schlichte Tatsache kann also nicht eine besondere Form der Gesellschaft begründen. Den informationellen Teil des gesellschaftlichen Zusammenhanges jedoch zumindest teilweise und auf entscheidende Weise an technische Einrichtungen zu delegieren und aus solcher Delegation neuartige Verhältnisse zwischen Menschen zu begründen – was ja zweifellos der Fall ist –, dies kann eine Formation von Gesellschaft begründen. Jedoch ist es nicht die Information selbst, die konstituierend wirkt, sondern die Technik zu ihrer Verarbeitung, die Informationstechnik also. Denn in dieser Tatsache, in der technischen Bearbeitung von Information, nicht aber in der üblichen, schlicht in allen Gesellschaften gegebenen Bedeutung von Information, liegt die spezifische Differenz der Informationsgesellschaft. Was diese technische Bearbeitung im einzelnen heißen mag, ist vielfach der Literatur zu entnehmen und nicht unser Gegenstand.

Der merkwürdige aktuelle Befund lautet also, daß, was als Synthesis der gesellschaftlichen Formation identifiziert wird, die Information, diese Rolle gar nicht spielen kann. Zumindest im Begriff ist die Informationsgesellschaft von vornherein ideologisch geprägt, indem sie *Information* sagt, wo es *Informationstechnik* heißen müßte.

Mit der *Informationstechnik* als entscheidender und treibender Tatsache bliebe die Entwicklung der bürgerlichen Gesellschaft viel stärker ihrer Tradition verhaftet, in der die Entwicklung der Technik auf wissenschaftlicher Basis stets

¹¹ ..., die wenig mehr für sich beanspruchen kann, als daß Maschinenmodelle des Menschen stets neuester Technologie folgen.

¹² Das englische Wort *memory* wird unglücklicherweise für beides verwendet, das menschliche Gedächtnis und den Computerspeicher. Dies wird gewiß zur verwirrenden Gleichsetzung beitragen, die nicht sein müßte.

¹³ Alfred Sohn-Rethel (1970) benutzt diesen Begriff, um die konstituierenden Merkmale einer Gesellschaft zu erfassen.

eine wesentliche Rolle spielte. Seit Descartes wird hier veräußert und mathematisiert, was in anderen Zeiten und Gegenden der Welt geheiligt wird.

Zeichen, Signale, Daten

Für das Weitere ist es notwendig, kurz die grundlegenden Begriffe der Semiotik zu rekapitulieren. Sich in grundlegenden Fragen der Informatik und Informationstechnik auf die Semiotik zu stützen, ist kein Akt der Willkür, sondern geradezu zwangsläufig. Denn die Maschinisierung der Kopfarbeit führt mit Notwendigkeit zu Gegenständen semiotischer Natur.

Die Semiotik wurde als allgemeine Theorie der Zeichen vor allem von Charles Sanders Peirce geprägt, dessen Werk erst neuerdings in seiner Bedeutung für die Postmoderne wirklich erfaßt wird. Die oft schwer zu durchschauenden und scheinbar widersprüchlichen Schriften von Peirce erfahren derzeit eine Wieder- und Neuauflage. Ich stütze mich auf (Peirce 1983, Walther 1974).

Der wichtigste begriffliche Schritt ist es, das Zeichen nicht als Ding, sondern als Relation einzuführen. Diese Relation wird dreistellig angenommen, indem das Zeichen als ein Erstes aufgefaßt wird, das für ein Zweites steht und (einem Menschen) ein Drittes bedeutet. Das Erste ist das Mittel des Zeichen, das *Repräsentamen*. Das Zweite ist das im Zeichen bezeichnete *Objekt*. Das Dritte ist die im beobachtenden Subjekt durch das Eintreffen des Zeichenereignisses hervorgerufene Bedeutung. Soll sie selbst wieder benannt werden, so muß sie explizit gemacht werden, was im *Interpretanten* geschieht.

Das Zeichen ist also als ein Tripel (Repräsentamen, Objekt, Interpretant) darstellbar. Der Interpretant ist selbst wieder ein Zeichen, denn ich kann die Bedeutung, die ich dem Zeichenereignis zuschreibe, nur durch ein weiteres Zeichen ausdrücken. Der Peircesche Begriff des Zeichens ist also subjektiv und in sich rekursiv. Gerade das macht ihn für heutige Analysen und für die Informatik so wichtig.

Das Objekt des Zeichens ist jenes Ding, jener Prozeß, jenes Ereignis, jene Vorstellung, die quer zu einer gesellschaftlichen Gruppe, in einer Kultur, insofern also intersubjektiv, dem Zeichenmittel zugeordnet ist. Es ist das, was das Zeichen bezeichnet, soweit als möglich unabhängig von der subjektiven Bedeutung, die ihm durch dieses oder jenes Individuum beigemessen wird.

Im Repräsentamen haben wir das Zeichen am deutlichsten nur als Ding, als stoffliches Substrat, als sinnlich wahrnehmbaren Gegenstand. Was es auch sei, Klang, Bild, Gedrucktes, Geste, Bewegung, das Repräsentamen ist sinnlich wahrnehmbar, damit stofflich vorhanden und deswegen wird es oft mit dem Zeichen selbst verwechselt. Wir betonen noch einmal, daß das Zeichen eine Relation ist, in die das Repräsentamen als *eine* Komponente, als die augenfällige Komponente allerdings, eingeht.

Ein Subjekt ist für die Zeichenereignisse notwendigerweise als vorhanden und beteiligt anzusehen. Ohne das Subjekt kommt es nicht zum Vorgang der Interpretation, zur Zuschreibung von Bedeutung. Das Subjekt bleibt aber außerhalb des Zeichenbegriffes, es ist nicht Bestandteil der Theorie selbst. Diese Semiotik ist keine psychologische Theorie.

Wir betrachten Zeichen syntaktisch, semantisch und pragmatisch. Die syntaktische Betrachtung fragt "Wie wird bezeichnet?" und hat es mit den Repräsentanten der Zeichen allein zu tun, mit der Selektion und Komposition von Zeichenmitteln. Die semantische Betrachtung fragt "Was wird bezeichnet?" und hat es mit dem Verhältnis der Repräsentanten zu den Objekten zu tun, mit der Signifikation. Die pragmatische Betrachtung fragt "Wozu wird bezeichnet?" und hat es mit den Interpretanten der Zeichen zu tun, mit der Interpretation.

Wir ordnen nun der syntaktischen Dimension *Daten* zu, der semantischen Dimension *Information*, der pragmatischen Dimension *Wissen*. Diese Zuordnung ist nicht so zu verstehen, daß die jeweilige Dimension das zugeordnete Verhältnis umfaßte, beinhaltete, erklärte. Sie ist vielmehr so zu verstehen, daß es in der syntaktischen Dimension zur Erscheinung der Daten kommt, in der semantischen zu der der Information, in der pragmatischen zu der des Wissens. Das ist noch genauer darzulegen.

Wir unterstellen hier eine Auffassung von *Wissen* als die Gesamtheit dessen, was ein Individuum (als autonomes System in Kooperation und Kommunikation mit anderen autonomen Systemen) dazu befähigt, in seiner Umwelt zu bestehen und sich zu entwickeln. Zu bestehen heißt nichts anderes, als sich zu entwickeln. Denn es gibt kein Überleben außer im ständigen Wechselspiel mit der Umwelt und dieses Wechselspiel führt zu einem ständigen Anpassen der inneren Strukturen an die sich ständig ändernden äußeren Bedingungen. Bekanntlich nennen wir diesen Vorgang *Lernen*. (Maturana & Varela 1987)

Alles das also zusammengekommen, was das Überleben eines autonomen Systems ermöglicht, sei dessen *Wissen* genannt. Nur in seinen niedrigsten Verästelungen kann diesem Wissen die äußere Form von Sätzen oder anderen Aufzeichnungen gegeben werden. Es besitzt nicht von vornherein explizite Form. Diese kann ihm nur auf besondere Veranlassung gegeben werden. Wissen ist nach dieser Auffassung nicht nur ausnahmsweise und gelegentlich implizit. Es ist vielmehr seinem *Wesen nach* implizit. Was dann, ausgehend vom impliziten Wissen expliziert werden kann, ist selbst nicht mehr Wissen, da aus dem autonomen System entfernt. Der explizit gemachte Teil eines Wissens ist vielmehr ein anderes, eben das Ergebnis des Versuches, mitzuteilen (also zu teilen!) und das nach außen zu kehren, was das Überleben des Systems sichern hilft.

Wir unterstellen weiterhin eine Auffassung von *Information*, die das erfaßt, was den Mitgliedern einer Gruppierung oder einer sozialen Schicht oder einer Kultur gemeinsam zukommt und wert und wichtig ist hinsichtlich der Bezeichnung von Dingen, Prozessen und Ereignissen in ihrem sozialen Zusammenhang. Soweit es im Bereich von Zeichen (also Kommunikation) überhaupt Objektives geben mag, ist es in der Information vorhanden. Information erweist sich bei dieser Auffassung als der quasi organisationale Teil des Wissens, als das, was den Mitgliedern einer Gesellschaft als *gemeinsames* Wissen erscheint. Information ist ihr Mittel der Verständigung. Auf soziale Information schrumpft, soll es zur Kommunikation kommen, individuelles Wissen. Information ist der intersubjektive Teil des Wissens, die objektivierte Annäherung an das subjektive Wissen.

Gehen wir schließlich in der Objektivierung noch einen Schritt weiter, fragen nach dem, was von der Information, die immer noch eine (zweistellige) Relation ist, übrig bleibt, wenn wir die Relation auf die eine ihrer dinglichen Seiten reduzieren, so gelangen wir zu *Daten*. Daten sind die bezeichnenden Mittel. Sie können, wo wir das wollen, aufgelistet und katalogisiert werden. Weil wir das nie tun, ohne Erklärungen dazu abzugeben, betten wir Daten gleich wieder in tendenziell bedeutungsvolle Kontexte ein. Die Daten erscheinen uns deswegen immer schon als Bedeutungen, sind es aber oder haben sie nicht. Doch wir spüren, daß wir mit Daten eine noch höhere Verbindlichkeit erreichen als mit der Information. Genauer gesagt, erreichen wir in den Daten, das Wissen betreffend, gerade *keine* Verbindlichkeit, sondern eine Unverbundenheit. Die vermeintliche Verbindlichkeit resultiert wohl daher, daß wir im Hintergrund die Maschine schon ahnen und ihr ganz zu Recht, wenn auch weitgehend unbewußt, unterstellen, sie ginge auf sinnvolle, verbindliche Weise mit Daten um, wenngleich wir doch wissen, daß Kontext und Situation ihr nichts sind. Kontext und Situation können, ja: dürfen!, der Maschine nichts sein, da wir sie doch gerade so einrichten wollen, daß sie stets und präzise und in großer Geschwindigkeit so kalkuliert, wie wir das wollen.

Nicht Information und Zeichen, sondern Daten und Signale

Wir betrachten das am Ende des vorausgegangenen Abschnitts Gesagte nun noch ein wenig genauer. Um Information per Computer zu bearbeiten, ist sie von ihren Kontexten zu befreien, und zwar von *allen* ihren Kontexten bis auf einen einzigen: den der Unterscheidung. Nur noch das eine vom anderen als ein anderes, als verschieden, trennen zu können, macht Daten aus¹⁴. Selbstredend bilden wir aus rudimentären, einfachen Daten sofort Strukturen und Komplexe von Daten, die das Bild ein wenig anreichern. Das Prinzip jedoch bleibt bestehen, daß dann nämlich, wenn wir von allen Zusammenhängen absehen, kurz bevor wir auf die Leerformel des *Einen und Insgesamen*¹⁵ zurückfallen, als einzige Qualität noch die Unterscheidung und Verschiedenartigkeit übrig bleibt. Sie ermöglicht uns, zu zählen und also zu rechnen, also mit Quantitäten umzugehen. Rechnen erweist sich als maximal dekontextualisierte Form des Denkens.

Rechnen kann nun auch ein Computer, wie er schon durch seinen Namen verrät. Und entgegen noch immer anzutreffenden Trivialglaubens¹⁶ kann ein Computer auch nicht mehr als das, nichts anderes als Rechnen. Er rechnet und er rechnet nach einer Beschreibung, die man bekanntlich Programm¹⁷ nennt und die sehr komplex sein kann. Das Erstaunliche an dieser trivialen Feststellung ist ja nicht, daß der Computer *nur rechnen* kann, sondern: daß *sovieles berechenbar* beschrieben werden kann, daß also *nur zu rechnen* heißt: *weit auszugreifen*. So als ob der kurze Griff des Rechnens einen langen Arm besäße.

Ohne Frage ist Rechnen eine *geistige* Tätigkeit, eine Tätigkeit also, die Intelligenz verlangt. Weil das so ist, sind die Leute nicht einfach dumm, die sagen, ein Computer könne denken. Er kann denken, nämlich auf die trivialste Weise, die wir kennen: maximal dekontextualisiert. Indem er die einfachste Form geistiger Tätigkeit in funktionaler Äquivalenz eingepflanzt erhält, kann er aber gerade nicht *denken* in einem umfassenderen Sinne des Wortes, sondern genau nur jene eine und einzige Art geistiger Tätigkeit funktional modellieren. Das gerade, was uns das Denken wichtig macht, die Beachtung wechselnder und immer neuer Zusammenhänge, das macht es dem Computer unmöglich zu denken. Das menschliche Denken ist soweit gekommen, daß es *eine* Form des Denkens (eben das Rechnen) so sehr von den verschlungenen, undurchschaubaren Weisen befreien konnte, in denen jeder einzelne Mensch es praktisch tut, daß das allgemeine Schema all jener individuellen Rechenweisen gefunden und sogar an Maschinen delegiert werden konnte. Klar, daß es dabei mit Notwendigkeit zu Standardisierungen kommen mußte.

¹⁴ Hierin erscheint der weitgehend nicht-relationale Charakter der Daten.

¹⁵ Im religiösen Zusammenhang heißt das "Gott". Er ist das inhaltsloseste Zeichen, das Menschen schaffen. Ein hoher Trost.

¹⁶ Ohne darauf näher einzugehen, spiele ich hier auf die lustigen Versuche an, künstlich Intelligenz zu schaffen, ohne lebendige Körper zu erzeugen.

¹⁷ Programme sind endliche Beschreibungen unendlicher Mengen von Prozessen.

Kontextfreie Information nun nennen wir Daten. Daten sind Zeichen, die auf ihre syntaktische Dimension reduziert worden sind¹⁸. Es gibt keine Computeranwendung, die nicht von Zeichenprozessen ausginge, denn Beschreibungen sind der Stoff, den ein Computer akzeptieren kann (Nake 1997), Beschreibungen als Daten oder als Programm. Beschreibungen sind aber Zeichen: sie stellen zwischen dem Material, aus dem sie gemacht sind, und den Dingen oder Prozessen, die sie beschreiben, eine Relation her.

Wenn Zeichenprozesse also das sind, womit wir beginnen müssen, wenn wir den Computer verwenden, so werden diese im nächsten Augenblick aber als wirkliche Zeichenprozesse bereits ignoriert, indem die Zeichen durch maximale Dekontextualisierung auf Signale reduziert werden (Nake 1997). Signale sind dann zunächst nur die syntaktischen Komponenten von Zeichen. Wir können sie jedoch wieder als Zeichen konstituieren, indem wir ihnen auf triviale Weise Semantik und Pragmatik zuordnen. Das Signal steht für sich selbst (seine Semantik) und bedeutet auch es selbst (seine Pragmatik)¹⁹.

Wunderbarerweise bewirkt die auf Signale reduzierte Beschreibung dann aber in der algorithmischen Bearbeitung etwas, das dem "draußen" an der Schnittstelle lauernden Menschen sofort wieder als Zeichen im vollen Sinne des Begriffs erscheint, also mit Bedeutung versehen (Nake 1994). Denn der Mensch – und nur er – bettet das erscheinende Signal sofort in Kontexte ein. Er kann gar nicht anders.

An der Schnittstelle zum Menschen bemerkt dieser bei der Ein- wie bei der Ausgabe nur Zeichen. Diese nimmt er als Information wahr. Der Durchgang durch die Schnittstelle wirkt wie eine Spur-Änderung: das Zeichen wird zum Signal, die Information wird zum Datum. Drinnen, in der Maschine wird das Signal bearbeitet und weitergereicht, ist Information auf Daten reduziert. Die offene, unendlich fortsetzbare Interpretierbarkeit des Zeichens weicht der eindeutigen Determiniertheit des Signals. Wir wollen das auch so haben, da wir nicht zulassen, daß die Maschine etwas macht, das wir nicht so erwarten.

Haben wir also zunächst gesehen, daß die Informationsgesellschaft nicht durch Information, sondern durch die Technik zu deren Bearbeitung bestimmt wird, so zeigt die nähere Betrachtung, daß der Computer als Kern dieser Technik gar nicht Information bearbeitet, sondern Daten. Die Rede von der Informationsgesellschaft, soweit sie die computerisierte Infrastruktur voraussetzt, erliegt einer zweifachen Mystifikation bzw. schafft diese.

¹⁸ Wir müssen etwas vorsichtiger sagen: nachdem Zeichen auf ihre syntaktische Dimension reduziert worden, also aus der semiotischen Welt herausgefallen sind, können sie auf eine triviale Weise wieder zu Zeichen gemacht werden. Dazu gleich mehr.

¹⁹ Dies entspricht auf semiotischer Ebene auch gerade dem, was wir vom Funktionieren von Programmen erwarten. In der Gleichheit von Repräsentamen, Objekt und Interpretant drückt sich die strenge Determiniertheit aus, die Software auszeichnen muß. Der Programmierer legt die eine einzige Bedeutung fest, die erwartungsgemäß der Prozessor dann auch hinein"interpretiert". Ahnen wir, wieso manche in Computer menschliche Fähigkeiten hineingeheimnissen?

Instrumentales Medium und semiotische Maschine

Wir neigen gern dazu, etwas uns neu Erscheinendes als das absolut Neue zu begreifen. Nachdenkend stellen wir dann fest, daß es immer schon Vorläufer gab, ja, daß das Neue seine Vorläufer produziert. Die Informationsgesellschaft – die wir nun halt trotz der eben geführten Begriffskritik so nennen wollen, Wittgenstein hat ja Recht – kann auf nichts anderem beruhen als darauf, daß ihren Produkt- und Waren-Strömen Doppel beigezelt werden, informationelle Doppel, Zeichenprozesse. Diese Zeichenprozesse laufen selbst mit hoher Geschwindigkeit ab. Sie bewirken, auf die Produktprozesse zurückgespiegelt, Beschleunigungen und Abkürzungen. Viele Transportvorgänge können tendenziell vereinfacht, beschleunigt, verknapppt werden, andere allerdings entstehen auch neu oder werden gesteigert. In den *Grundrissen* schreibt Marx:

Je mehr die Produktion auf dem Tauschwert, daher auf dem Austausch beruht, desto wichtiger werden für sie die physischen Bedingungen des Austauschs – Kommunikations- und Transportmittel. Das Kapital treibt seiner Natur nach über jede räumliche Schranke hinaus. Die Schöpfung der physischen Bedingungen des Austauschs – von Kommunikations- und Transportmitteln – wird also für es in ganz anderem Maße zur Notwendigkeit – die Vernichtung des Raumes durch die Zeit. (Marx 1971, 423)

Die Kommunikationsmittel werden für das Kapital, je mehr es sich global entfaltet, umso bedeutsamer. Es kann uns bei dieser Vorhersage Marx' nicht darauf ankommen, ob er eine konkret technische Form im Auge hatte oder nicht. Wichtig ist die generelle ökonomische Funktion, die er anspricht. Das dem Kapital innewohnende Moment – die Überwindung aller räumlichen Schranken – erleben wir mit dem Internet in einer besonderen, fortgeschrittenen Form. Sie konnte, ihrer Natur als allgemeine Infrastruktur gemäß, nicht vom Einzelkapital geschaffen werden, wenn auch einzelne Kapitale in der Vorbereitung besondere Bedeutung gehabt haben mögen. Sie ist vielmehr vom Gesamtkapital zu schaffen und zu organisieren. Ihre Wirkung ist die Vernichtung des Raumes durch die Zeit – kann man knapper und klarer ausdrücken, was heute Scharen von Forschern bewegt?

Zu der Zeit, da die Grundlagen des Rechnens in aller Strenge verstanden werden mußten und Alan Turing das theoretische Modell dafür schuf, entstand auch schon fast im Gleichklang die Maschine, die das übernehmen konnte. Bis der Computer als semiotische Maschine gekennzeichnet werden konnte (Nadin, Coy), dauerte es zwar noch einige Jahre. Doch dann wurde auch bald klar, daß wir es wegen des semiotischen Charakters der Abläufe auf dieser Maschine und der Gegenstände, die sie bearbeitet, mit einem Medium zu tun haben. Einem Medium nun zwar, daß durch seine inhärent instrumentalen Besonderheiten weiter zu kennzeichnen ist. Heidi Schelhowe hat das mit dem Begriff des *instrumentalen Mediums* getan (Schelhowe 1997).

Der Computer hat auf Grund seiner beiden wichtigsten Komponenten, Speicher und Prozessor, sowohl verarbeitende wie bewahrende Funktion. Er vermittelt zwischen Menschen (kommunikative Funktion) und er vermittelt zwischen

Mensch und Objekt (instrumentale Funktion). Die Steigerung der Kommunikationsmittel, von der Marx spricht, geht einher mit einem Aufsaugen instrumentaler Funktionen und einer allgemeinen kulturellen Wirkung.

Wenn dies auch alles wenig mit der manchmal mißverstandenen Universalität des Computers zu tun hat, so ist sein Einwirkungsfeld, ohne daß man zwischenzeitlich einen instrumentalen Wechsel vornehmen müßte, doch gewaltig. Es wird verständlich, warum es in journalistischen, populären Aufsätzen immer wieder zur Überhöhung dieser Maschine kommt.

Dennoch! Wir leben in dem notwendig falschen Bewußtsein von einer Gesellschaft, der es immer mehr und konzentrierter um Wissen ginge, in der Wissen aber nicht zur ruhigen Weisheit führt, sondern zum aufgeregten Klicken verführt, weil das, was da scheinbar schwerelos bewegt wird, Daten sind und keine Information. Eine ganze Gesellschaft kann ein notwendig falsches Bewußtsein von sich besitzen und dennoch einigermaßen zurechtkommen. Die Widersprüche werden spitzer werden. Doch sie werden lange aushaltbar bleiben.

Eine günstige Wendung mag dennoch sein, die maschinelle Bearbeitung von Information stets und überall in ihren Kontexten zu halten, ja: die Kontexte bewußt zu betonen, da sie in der technischen Konzentration spontan vernachlässigt werden. In einem Projekt über die Verwendung von Computern in der Grundschule in Bremen schälte sich rasch sehr deutlich heraus, daß der abschließende Charakter der Technik nach der Form des Offenen Unterrichtes verlangte. Der Kontext (Schulunterricht) setzte sich gegenüber der Technik durch.

In einem anderen Fall entwickelten wir zwei Hypermedien als permanente Installationen für das Stadtmuseum Delmenhorst. Dabei konnten wir erfahren, daß die gezielte Aufnahme der Besonderheiten des Museums-Ambientes sich in vielfacher Weise günstig auf die Gestaltung des Mediums auswirkte. Eine CD-ROM haben wir nicht produziert. Denn die Installationen sind nicht in erster Linie Software, der man sich unabhängig von Ort und Zeit aussetzen kann. Sie stellen vielmehr eine Veränderung des Museums-Ambientes dar und müssen in diesem auch verwendet werden. Die Zeichenwelt, die uns in diesen Fällen über den Bildschirm im Rahmen des Museums zugänglich gemacht wird, gewinnt ihre semantische und pragmatische Sättigung durch eben diesen Rahmen.

Wir ahnen, daß die Informationsgesellschaft es nicht mit der rasenden Übermittlung und Verarbeitung von Informationen (oder Daten) zu tun hat, sondern vielmehr mit der Gestaltung von Medien als (sozialen) Prozessen (Wilkens & Nake 1998). Doch das ist ein weiteres, neues Feld.

Literatur

[Barrett & Redmond 1995]

Edward Barrett, Marie Redmond (eds.): *Contextual media. Multimedia and interpretation*. Cambridge, MA: MIT Press 1995 (pbk. 1997)

[Eisenstein 1983]

E. Eisenstein: *The printing revolution in early modern Europe*. Cambridge, MA: Cambridge University Press 1983

[Foerster et al. o.J.]

Heinz von Foerster, Ernst von Glasersfeld, Peter M. Hejl: *Einführung in den Konstruktivismus*. München: Piper

[Glasersfeld 1997]

Ernst von Glasersfeld: *Radikaler Konstruktivismus. Ideen, Ergebnisse, Probleme*. Frankfurt/Main: Suhrkamp 1997

[Marx 1971]

Karl Marx: *Grundrisse. Ergebnisse des Produktionsprozesses*. Frankfurt: Europäische Verlagsanstalt 1971

[Maturana & Varela 1987]

Humberto R. Maturana, Francisco J. Varela: *Der Baum der Erkenntnis*. Bern, München, Wien: Scherz 1987

[Nake 1992]

Frieder Nake: Informatik und die Maschinisierung von Kopfarbeit. In: W. Coy, F. Nake, J.M. Pflüger, A. Rolf, J. Seetzen, D. Siefkes, R. Stransfeld (Hrsg.): *Sichtweisen der Informatik*. Braunschweig: Vieweg 1992, 181-201

[Nake 1994]

Frieder Nake: Human-computer interaction: signs and signals interfacing. *Languages of Design 2* (1994) 193-205

[Nake 1997]

Frieder Nake: Der semiotische Charakter der informatischen Gegenstände. In: *signum um signum*. Elisabeth Walther-Bense zu Ehren. *Semiosis Heft* 85-90 (1997) 24-35

[Negroponte 1995]

Nicholas Negroponte: *Being digital*. New York: Alfred A. Knopf 1995

[Peirce 1983]

Charles S. Peirce: *Phänomen und Logik der Zeichen*. Herausgeg. u. übersetzt von Helmut Pape. Frankfurt/Main: Suhrkamp 1983

[Postman 1993]

Neil Postman: *Technopoly. The surrender of culture to technology*. New York: Alfred A. Knopf 1993

[Schacter 1996]

Daniel L. Schacter: *Searching for memory. The brain, the mind, and the past*. New York: Basic Books 1996

[Schelhowe 1997]

Heidi Schelhowe: *Das Medium aus der Maschine*. Frankfurt, New York: Campus 1997

[Shannon & Weaver 1948]

Claude Shannon, Warren Weaver: The mathematical theory of communication. *Bell System Technical Journal* July 1948

[Sohn-Rethel 1970]

Alfred Sohn-Rethel: *Geistige und körperliche Arbeit*. Frankfurt: Suhrkamp. (In überarbeiteter Neuauflage 1989 bei VCH, Weinheim)

[Spinner 1998]

Helmut F. Spinner: *Die Architektur der Informationsgesellschaft*. Bodenheim: Philo Verlagsgesellschaft 1998

[Stoll 1995]

Clifford Stoll: *Silicon snake oil. Second thoughts on the information highway*. New York: Doubleday (Anchor Books) 1995

[Varela 1993]

Francisco Varela: *Kognitionswissenschaft, Kognitionstechnik*. Frankfurt/Main: Suhrkamp 1993

[Walther 1974]

Elisabeth Walther: *Allgemeine Zeichenlehre. Einführung in die Grundlagen der Semiotik*; Stuttgart: Deutsche Verlagsanstalt 1974

[Wilkens & Nake 1998]

Ulrike Wilkens, Frieder Nake: *Das Medium als Prozeß. Eine virtuelle Burg im Stadtmuseum Delmenhorst. FIFF Kommunikation* 3/98 (Sept. 1998) 50-56

Information Problems in Information Society

Antonín Rosický

Department of Systems Analysis,
University of Economics, Prague
130 00 Praha 3, Churchill sq. 4, Czech Republic,
e-mail: Rosicky@VSE.CZ

... difficulties arise from a theoretical and methodological design framework that rests upon an underconceptualised notion of information.

... in living systems in general, and social systems in particular, evolution and information are so tightly interrelated that it is not appropriate to discuss one without the other.

Bela Banathy (1996)

1. Introduction

Our tendency to name important period of human history is not only a vagary, it briefly characterises main features of an adequate society (or social systems) and its development. This name is a metaphor that help us to better understand this society and affects a perception of its general problems and ways of their solution too.

However the used term cannot comprehend elusive changeable and very intricate nature of human systems. Moreover each metaphor expresses the reality only from partial view and it could be misleading if it is not well grasped. Big idea of the information society as well as hopes and doubts connected with it are results of similar and conceptual obscurities.

2. New (information) society

Thirty years ago probably most from us believed that the name "century of the atom" is fitting to the coming era and better future of human. Today many apprehends of the nuclear world and puts expectancy on information or more precisely on information technology.

Just many important arguments advocated the entrance into information society arises from famous Porat's studies based on empirical observation of some social and economics facts in United States in sixties and seventies. Nevertheless the later (in eighties) this idea was revised: Nobody today doubts

about the fundamental importance of information in society and its future development, but many authors points to other faces of newly emergent society. For example:

- Technocratic Society (Zbigniew Brzezinski)
- Post-industrial Society (Daniel Bell)
- Post-capitalist Society (Peter Drucker)
- Post-fordist Society (Lucas Introna)

Each of them emphasizes some different attributes of recently arising society and most of them connect its character with influenced significance of information or knowledge. In spite of the fact that the importance of information in many societal domains has been mentioned (Hayek 1945) and theoretical issues was founded in cybernetic (Wiener, 1948), a common awareness of it is connected with a development of computers and "digitally" oriented information technology. Despite remarkable success of information technology use many problems in society are not solved, other manifest more sharply and even furthers recently arise.

Undoubtedly each stage of the human society has specific problems, however present challenges might be distinguished by few traits, let us stress only two from them:

- They present a global character, at the same time they arise and influence events around the whole world. As such they causes "turbulent character" of the world in which the uncertainty is a natural constituent of its order.
- Firstly in history they present — thanks to capabilities of the modern technology — the "global danger": the possibility to change, to break, to jeopardize and even to destroy the life on the Earth. Probably many apparent threats (for example an atom bomb) are less exposure than many other hidden matters (arising for example from genetic engineering).

Reasonably only a few reflecting the recent character of the society don't consider about information and don't point out at the increasing significance of knowledge and information. Actually the essential changes characterizing a movement to information society originate from increasing complexity of the "global world" that is narrowly connected just with new possibilities to handle information. However recognized importance of information within all domains of our life including politics, business, education and others social domains is not attended by a better understanding of information. And the use of new and unprecedented possibilities to treat information through "old ways of thinking" cannot provide expected benefits.

3. Phenomena of the information

A common comprehension of information arises from a mechanical and objective principles of positivistic thinking. In this sense the information is a strongly given entity reflecting (mirroring) or representing the reality. As such the amount of information decreases uncertainty and allows to decide and to plan in the better way.

Probably most important problem in this "information confusion" is sweeping of divers aspects or types of information that emerge along with in-

creasing complexity of considered systems. Regarding the term "information" we do not distinguish the different nature of the mentioned concept abundantly and consequently; we confuse and simplify it in an unacceptably way. As a result of that clutter we manage information — we substitute dealing with information for a data manipulation excepting the reduction of uncertainty...

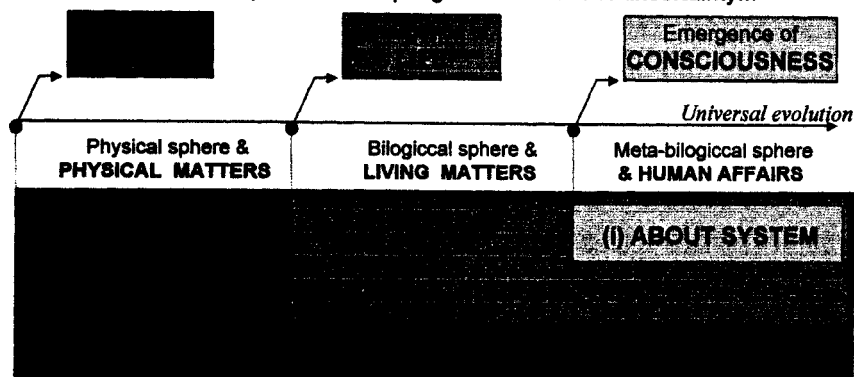


Figure 1 Three emergent aspects of information, depended on adequate properties of the universe (type of system)

For the next consideration we can use categories of nature, suggested by Jonas Salk (1983) who defines three main eras of the universal evolution and increasing complexity. He outlines three types of systems (or matter): physical, biological, each of them is characterised by new emergent and essential properties. The emergence of new system's property narrowly incident to new aspects (or types) of information. For systems on each higher level are important aspects of information from lower levels, however new aspect emerges and changes the concept of the system's complexity - the situation is brief and well arranged in the figure 1.

On the lowest level of physical systems the ("physical") information is an intrinsic property of systems. It is (internal) **information** in system that has an objective character and determines its order and/or structure on the one side and characterizes its state on the other.

The second (biological) sphere introduces new quality - living systems, that are autonomous towards their environment and adapting to it. Internal (i.e. genetic) information of living systems is transmitted from one generation to the next in replication process with random character of "nature selection".

However newly emerging aspect (type, concept) of information emerges on this level and plays essential role for adaptation of living system which receives **information** from its environment. Two important facts implicate its essential properties:

- The environment presents higher (wider) system with internal information. However received information depends on properties of receptors and particular condition too - it has not objective (but nor subjective) character.

- Such information is impetus (rather than input) that initiates certain response of the receiver - his behaviour. Important condition necessary for systems evolution (adaptation of species) is the possibility of different ("individual") responses to the same received information - it doesn't determine response strongly but receiving system has certain looseness (information is more loose) in its responding action.

The relations between receiving information and resulting behaviour depend upon the receiver, respectively on a complexity and quality of his "nervous systems". The more complex nervous systems use information from last experiences, they are capable to learn himself and use some kind of "knowledge". As well as the experiences result from an individual interaction with the environment (higher system) mentioned relations (and knowledge) have individual character. From a general view the resulting behaviour becomes more loose similarly as a "meaning" of considered information.

The highest systems level results from the emergence of human conscious, and mind's activities that are often considered as information handling. The conscious gives an unique and essential possibility of abstract thinking and perform on symbols. Bertalanffy (1972) stresses this essential attribute of human calling him "*animal symbolicus*". Human - thanks to his conscious - aware himself in contrary of the other world - produce **information about system**.

This information is presented through symbols which form is commonly named data¹). Not before this level we can consider about "human" information (Stonier, 1990) and semiotic character (no only comprehension) of the information (Stamper, 1973) obtains important role. Meaning of shared information is interpreted during human mental processes in that individual (personal) knowledge plays fundamental role.

Let us notice a few considerable aspects of this type of information:

- Human, symbolic presented information, is a product of individual mind. It presents a result of human cognition and thinking including individual intentionality (Searle, 1986) as individual opinions, intentions (inc. to desinform), believes.
- This information expresses a vision of system together with a relationship of its "author" (originator) to this system. However it has not "objective constitution" it is a part of an objectively existing information in social systems.
- As a received information has a character of mentioned impetus, influencing and triggering behavior of other systems elements - people, animals and machines. This behavior is determined in the case of machines and less or more loose in the case of people.

¹ > Many authors accept the fact that data present a form of information only with hesitation. They confuse the question "What are data ?" with the question "What data express ?". In addition to that they confound the answer too and reduce it upon "The evident or better measured and quantified facts."

- Different ways of the same data shared in system result into different meanings — shared information cover certain effect of uncertainty (mostly beyond description in terms of probability).
- This type information needs not to decrease uncertainty, very often brings obscurity, doubtfulness and uncertainty. The issue of its validity and credibility (rather than "truthfulness") secures essential significance.
- To decrease a danger of misinterpretation and understanding within systems needs better communication of information as a shared of its meaning.

It is not possible to degrade the communication of (human) information on transmission of data (symbols).

4. Tangle of information and knowledge

While the distinction between information and data is mentioned for a long time and it is covered in each better textbook the distinction between information and knowledge enjoys a higher attention in the recent time. Nevertheless their concepts are more complicated and confused - usually we depict information and knowledge in a mutual interaction.

Both concepts are interlaced in common using: Cambridge International Dictionary of English defines:

- ◆ *Information* is a knowledge about something, esp. facts and news.
- ◆ *Knowledge* (is) understanding of or information about a subject which has been obtained by experience or study, and which is either in person's mind or processed by people generally.

In this sense information is considered as the notion "information in system" mentioned above, while "knowledge" cover the concept of "information about system". The common notion of information covers (and confuses) two very important concepts:

- ◆ information as a "internal property of universe" ("negentropy", which determine system)
- ◆ information as a knowledge (that is result of individual human cognition).

Similarly many modern authors (Miller 1978; Stonier, 1990; Kampis, 1991; Banathy, 1997) consider more aspects of information, some from them speak about "informations" (Pedersen, 1996). Very interesting is Miller's work, elucidating different form or aspects of information within living systems and their relations and mutual changes. Also Kampis considers two aspects of information: In the sense of knowledge he names it as "nonreferential information" and in the different (opposite) sense he defines "referential information" and connects it with actions that results into changes in (received) systems. Just this Kapis concepts of "information as action" corresponds to mentioned type "information from system". To accept third concept (aspect) of information labelled by Banathy (1997) as "state referential", we can brief arrange table No. 1, that compares the used terms and briefly characterizes three mentioned aspects of information, its denotation and implication.

An interchange and a conversion of information types create closed loops or "information circle". In other words: Mentioned circle covers associations of

knowledge, information (as actions / impetus) and states of the world. Regarding three mentioned types of information we can show described concept of information(s) by the following figure No. 2. The figure shows complicated relations between information that determines each system, human perception & cognition on the one side and (human) actions managing this system. Schematically drawn circle is actually much more complicated in a consequence of rich and naturally uncertain communication (or information flows) within social systems.

Table 1 The characterize and compare of three main aspects of information

Common name	Systems concept	Kamips/Banathy	Character
NEGENTROPY	Information in system	State referential information	It determines system and /or characterizes its state
INFORMATION	Information from system	Nonreferential information	It is interchanged between living system and its environment
KNOWLEDGE	Information about system	Referential information	Product of experiences, mental processes and mind

Sketched basic rotation of information conceives basic framework of evolution (including traditional evolution of living systems). This circle has a character of nature loop — it is not possible to interrupt it, however it is practicable — with consequences upon evolution. This fact raises ordinarily unthought dilemma of information handling, including the use of modern information technology and building/using of designed information systems. Therefore both mentioned activities as well as many others, that present pragmatically oriented trends of modern "information" society, need good understanding of information.

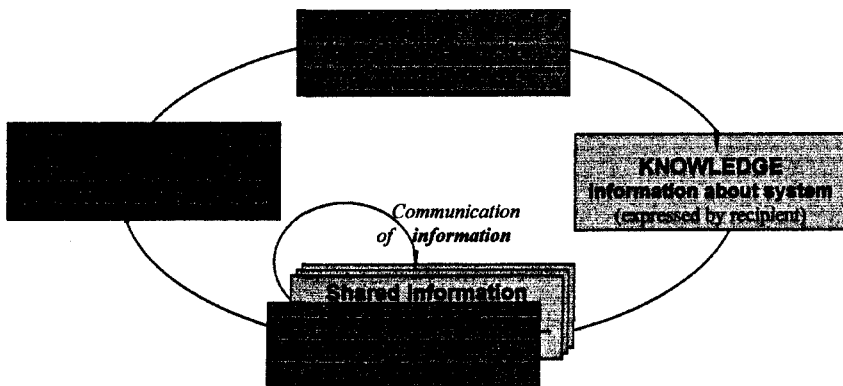


Figure 2 Closed information loop (circle): association of three aspects of information enriched about human communication of information in the society

However difficulties arising from getting over a mechanistic concept of information don't finish by acceptance of three proposed ideas of information.

Next from them arises the relative character of information: one is as well receiver of information (from system) as a producer of explicitly manifested revealed knowledge (information about system). Taking the consequences it becomes information from system for other potential receivers and changes internal information of system at the same time. Laws and enactment given from the to top public offers good examples: they determine social system, they influence behavior of knowing (receivers), they could be interpreted in different ways.

5 Entangling with data and information

More often mentioned and more typical is effort to render a nature of meaningful information and its confused relation to data and information. Also British Standard (BS 3527, Part 1, 1976) defines information as *"the meaning that human assigns to data by means of conventions used in their presentation."* Many authors suggest altered vision association between them, some try to amplify it just for a knowledge and yet another use some further entities as "wisdom" (Ackoff, 1989) or "capta" (= captured data, Checkland, 1998). They attempt to connect data and information through human mind in this way, however linear and causal relations cannot catch the nature of information. Typical result presents a "linear chain" schematically drawn in the figure 3.



Figure 3 Line, showing relations between information aspects

The meaning is associated just with information (information from system / as action) and its constitution qualified to biological and mental processes or to self-production and autopoiesis (Mingers, 1995). Just appropriate ideas and theories overstepping mechanical paradigm give new attributes to information, knowledge and meaning too. They are not steady and firmly given as a "content" or "representation" but they result from activities of living systems. Kampis, 1991, p. 436 suggests:

... information in cell is not about something. That is, information is not passive, representational and established, but active, specificational, and productive.

Regarding this actualities we can sketch new and necessarily simplifying schema (see figure 4) that demonstrate more intricate relations within (closed loop of) information processes resulting meaningful actions. The learning process creating intrinsic human knowledge plays fundamental role like this which this tacit knowledge engages in the process of interpretation data into meaningful information (as action).

6 Ambiguity of knowledge

Let us emphasize the distinction between "learning" and more passive "teaching"²): Learning process presents active perceptive and cognitive activities originating from individual human interaction with environment. Individual intrinsic knowledge results from experiences acquired in this interaction. On the opposite side the teaching activities are based upon the communication of explicitly declared information about system (that has a character of knowledge too). As such it information is presented through language symbols and need individual intrinsic knowledge.

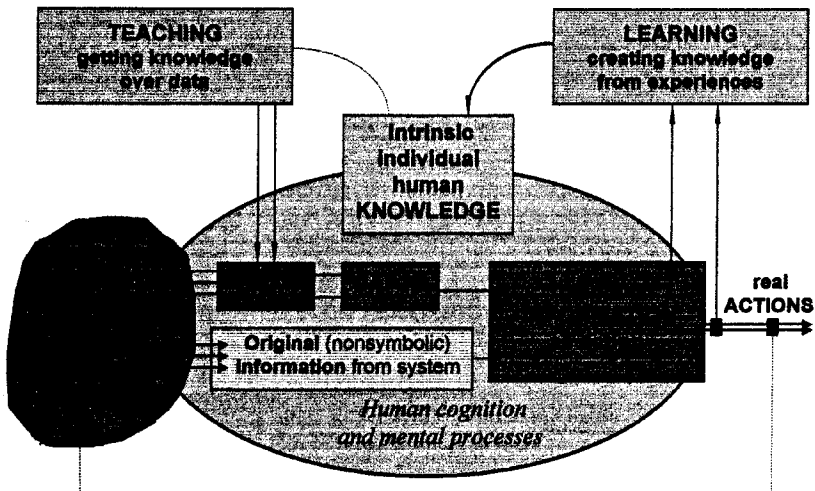


Figure 4 Information circle regarding human information handling

In spite of this intrinsic knowledge has a character of intimated general concept of "information about system" and constitutes a unique singleton. It is a important part of human fenophenotype and we can use a metaphor "it complements genetic information provision of the man" and puts the basis of his individuality. Intrinsic knowledge as well as cognition originates a nature of auto-poises, living systems and — on developed abstract level connecting with "symbolic function - of a human too. Regarding this a holistic character of intrinsic knowledge emerges and a comprehension of "knowledge (cognitive) system" appears more appropriate than thinking of the "single knowledge".

Let us point to a few important attributes of human intrinsic knowledge (system):

- It is not neither objective (independent upon the man) nor subjective (independent upon environment) — intrinsic knowledge is just individual (Wino-grad and Flores, 1986). Arising from unique interaction of the individual with

² > I am not quite sure if chosen terms distinguish intentioned ideas precisely: it is very difficult express so nice sense, moreover in a foreign language.

his environment in evolution process it covers some shared values, patterns of behavior, culture norms.

- It is based on abstract and symbolic cogitation — it has conceptual character closely associated with an ability to use languages and primarily based upon mother tongue. Coherence (no hierarchy) of syntactic, semantic and pragmatic aspects emanating from experiences enriches human knowledge with value, intentionality and probably with an emotionality too.
- Intrinsic knowledge appears as a dynamic and self-developing system. It is not based upon the relations named as representation, but it is "*actively built up by cognition, that serves the subject's organization of the experiential world, not the discovery of an objective ontological reality*" (von Glasersfeld, 1991, p.233).

It is necessary to differentiate carefully the intrinsic human knowledge from a meaning that is used in some other domains:

Firstly the concept of knowledge (and its representation) plays an essential role in the artificial intelligence primarily. In this case knowledge presents value free and formalized fact (factual knowledge) and resolutely defined rules for their using (procedural knowledge) handled independently from its user.

Similarly the concept of "organizational knowledge" and "organizational learning" is rather metaphor, however organization comprises resembling properties. Actually the intrinsic knowledge is the inborn attribute of biological single (person) and the adequate organizational feature emerges depending on information changes (see "information circle") within organization.

7. Conclusion for future information society

Just intrinsic knowledge constitutes fundamental framework of dealing with human information and consequently of an individual competence as ability to compete in evolution process. Individual character of intrinsic knowledge results into many different ways of interpreted information, its meanings and (free) individual behavior. Emergent uncertainty — undesirable from a mechanistic worldview — is fundamental attributes of successful (?) evolution of human species.

Better considering the nature of information and its role in evolution would contribute to better understanding of human nature as well as the problems of society, including more pragmatic activities from ecology and education, through business . up to the design and implementation of information systems or effectiveness of (modern) information technology.

Actually the importance of information in society doesn't increase — it has fundamental significance since long ago. What really increases are a necessity to survive in much more turbulent world of which complexity (amount of information - Kllr, 1991) originates from dramatic accelerating evolution in the last "industrial" era.

The huge amount of human information - information about systems, presented by data - have increased in an extraordinary way thank to many "artificial" activities and systems as scientific investigation, public education, mass-media. and information systems based on information technology (inc. internet

etc.). Understanding information we should agree with (Holtham's (1996) argumentation that *"We have so many different inputs of data now in most societies and organizations that it actually becomes harder to extract the useful information"*.

Our intrinsic knowledge - the framework for human information dealing - does not develop in an appropriate way. Any information technology cannot change it, however it can to support needed changes. In opposite: its fantastic power use in spirit of mechanical paradigm rather expands quantity of exploited data. Without the applicable receiver's (intrinsic) knowledge this situation baffles interpretation and causes information overload in this ways. Fundamental is our "paradigm shift".

From many particular ways let us point to the modification of the concept of "informatics" and its education (that commonly covers rather teaching than learning). Wider concept, including not only technologically oriented skills (manipulating with computers in worse case) other branches: systems associations and cognitive human abilities, is sketched on figure 5.



Figure 5 Wider concept of informatics - from a systems point of view

Each technology encloses certain amount of knowledge (Salomon, 1992), given by its authors: designer, architects, producers, organizers, engineers, initiators. but it is used by some others. Users are empowered thanks to use it without an adequate competence, without understanding of association and consequences. Natural interchange of information is cracked and the evolution process is effected and new issue of power emergent. Namely large information systems based on information technology, designed by many engineers produces little perceptible, anonymous — and therefore badly controlled — power.

By the way: Zbigniew Brzezinski (1993) — from entirely different positions - indicates two points that humanity must get in hand on the eve of new millennium: *genetic engineering and artificial intelligence*. Both cover information dealing.

Literature:

- Ackoff, R., [1989]: *From data to Wisdom*, in: Journal of Applied Systems Analysis, 19, pp. 3-10;
- Angell, I., O., [19975]: *Welcome to the "Brave New World"*, in: Information Systems: An Emerging Discipline? (eds. Mingers, J., Stowell F) , McGraw Hill, London;
- Banathy, B. [1997]: *Information, Evolution, and Change*, in: Systems Practice, Vol.10, No.1, p.59-83;
- von Bertalanffy L. [1967]: *Robots, Men and Minds*, Braziller, New York;
- Brzezinski, Z. [1993]: *Out of Control: Global Turmoil on the Even of 21st Century*, (Czech edition, Victoria Publishing, Praha, 1993);
- Ennals, R., [1991]: *Artificial Intelligence and Human Institutions*, Springer-Verlag, London`
- Von Glasersfeld, E. [1990]: *An Exposition of Constructivism: Why Some Like it Radical*, in: Klir, G. Facets of Systems Science, Pergamon Press, Oxford, 1991;
- Havel, V., [1984]: *Politics and Conscious*, in: Living in Truth (Vladislav, J. eds.), Faber and Faber Limitid, London, 1989;
- Hayek, F.A. [1945]: *The Use of Knowledge in Society*, In: The American Economic Review, Vol. 35, No. 4, pp. 519-530
- Holtham, C. [1996]: *Resolving the Imbalance between Information and Technology*, in: The Fourth Resource: Information and Its Management, ed. Best, D.P., Aslib Gower, Hampshire, 41/56;
- Checkland P., Holwell, S., [1998]: *Information, Systems and Information Systems*, Wiley, Chichester;
- Introna L., [1997]: *Management, Information and Power*, MacMillan, Hamshire and London, UK;
- Kampis, G., [1991]: *Self-modifying Systems in Biology and Cognitive Science*, Pergamon Press, Oxford;
- Klir, G. 1991: *Facets of Systems Science*, IFSR International Series on Science and Engineering, vol.7, Pergamon Press, Oxford;
- Liebenau, J. and Backhouse, J., [1990]: *Understanding Information*, London, MacMillan;
- Lorenz, K. [1983]: *Der Abbau des Menschlichen*, Piper, Mhnchen;
- Luhman, N. [1993]: *Ecological Communication: Copying with the Unknown*, in: Systems Practice, Vol. 6, No. 5, p. 527-539;
- Lyotard, J.,F. [1986]: *Le postmoderne expliquÇ aux enfants*, Editions GalilèÇ, France;
- Miller, J. [1990]: *Introduction : The Nature of Living system*, in. Behavioral Science, Vol. 35, No. 3;
- Mingers, J. [1995]: *Self-Producing Systems - Implication and Applications of Autopoises*, Plenum, New York;
- Pedersen, M.K. [1996]: *A Theory of Informations*, Samsundslitteratur Roskile Univeritetsforlag, Frederiksen (Denmark);
- Rosický, A. [1997]: *Information within the (Human) System*, in: Systems for Sustainability: People, Organizations and Environments, Plenum, New York;

- Rosický, A. [1998]: *The Danger of an Anonymous Power: (Human ?) Interface between computers and Business Organisation*, Proceedings of Interdisciplinary Information Management Talks 1998 (prepared paper);
- Salomon, J. [1992]: *Le Destin technologique*, Bailland, Paris;
- Searle J. [1986]: *Minds, Brains and Science*, Harvard Univ. Press, Massachusetts;
- Stamper, R. [1973]: *Information in Business and Administrative Systems*, Vatsford, London;
- Stonier, T. [1990]: *Information and the Internal Structure of the Universe*, Springer-Verlag, London;
- Walsham G. [1993]: *Interpreting Information Systems in Organizations*, J. Wiley, Chichester, UK;
- Weick, K. [1985]: *Cosmos vs. Chaos: Sense and Nonsense in Electronic Context*, in: Ruggless: *Knowledge Management Tools*, Butterworth-Heinemann, Oxford;
- Winograd, T. and Flores, F. [1986], *Understanding Computers and Cognition*, Ablex Pub. Norwood;

Informationsgesellschaft oder eine andere Gesellschaft?

Jiří Cejpek

Philosophische Fakultät der Karls-Universität Prag
Institut für Informationsstudien und Bibliothekswissenschaft
Celetná 20, 116 36 PRAHA 1, Tschechische Republik
e-mail: Jiri .Cejpek @ ff.cuni.cz

Inhalt

1. Einleitung
2. Begriff der Information als Problem
3. Vom Verständnis des Menschen zum Verständnis der Natur
4. Schlussbetrachtung

Zusammenfassung

In der Fachliteratur und Publizistik wird immer häufiger der Ausdruck „Informationsgesellschaft“ zur Bezeichnung der sog. entwickelten heutigen Gesellschaft benutzt. Die Kritik des Ausdrucks „Informationsgesellschaft“ geht von der Erkenntnis aus, dass aus der großen Menge der Bedeutungen des Informationsbegriffs am wichtigsten für den Menschen die Auffassung der Information als psychophysiologische Erscheinung und als Prozess ist. Die so aufgefasste Information ist jedoch durch die charakterlichen und moralischen Eigenschaften des Menschen bedingt, die für die weitere Entwicklung der Gesellschaft bestimmend sind. Zu der durch die Sinne des Menschen wahrgenommenen Umgebung gehören nicht nur das soziale Milieu, sondern auch das natürliche Umfeld (die Biosphäre). Deswegen schafft die Informationskommunikation die Voraussetzungen nicht nur für das zwischenmenschliche Verständnis und die Verständigung, sondern auch für die Überwindung des anwachsenden Konflikts zwischen der Ordnung der menschlichen Kultur und der Ordnung des Planeten Erde.

1. Einleitung

Die zukünftige Gesellschaft der sog. entwickelten Länder wird mit unterschiedlichen Attributen wie postindustriell, postmodern und auch als Informationsgesellschaft bezeichnet. Die Bezeichnung der Gesellschaft als postindustriell gibt an, dass die Bedeutung der industriellen Produktionsprozesse zurückgeht. Eine detailliertere Analyse dieses Ausdrucks zeigt dann, dass diese freigewordene Stelle von der Information besetzt wird, die als entscheidender Faktor für die Entwicklung der Gesellschaft aufgefasst wird. Das, was der Begriff „postindustrielle Gesellschaft“ eher andeutet, bringt ganz eindeutig der Terminus „Informationsgesellschaft“ zum Ausdruck.

Wahrscheinlich wurde der Ausdruck „Informationsgesellschaft“ zum erstenmal von A.L.Norman im Jahre 1975 benutzt. Er entstand infolge der sehr schnellen Entwicklung der Rechen- und Telekommunikationstechnik und -technologie und einer gewissen Erwartung deren Einflusses auf Mensch und Gesellschaft.

Heute verwendet man in verschiedenen Verbindungen sowohl die Ausdrücke „Information, Informations-, informatisch“ als auch „Informationsgesellschaft“, und das nicht nur in der Umgangssprache, sondern auch in zahlreichen Dokumenten solcher internationaler Organisationen wie z.B. UNO, UNESCO, EU, IFLA, FID usw.

Die immer mehr verbreitete Benutzung des Ausdrucks „Informationsgesellschaft“ löst eine Reihe Fragen aus, die nicht mehr nur terminologischen Charakter besitzen, sondern zum Wesen der sog. entwickelten Gesellschaft gerichtet sind. Werden wir wirklich zu einer Gesellschaft, für die die wichtigsten Angelegenheiten die Information, die Informations- und Kommunikationsprozesse sind? Ist dieser Terminus nicht einem Modell verpflichtet, ähnlich dem, der vor 20-30 Jahren den Terminus „Atomzeitalter“ geläufig machte? Und um welche Bedeutung des polysemantischen Begriffs Information geht es in der Verbindung „Informationsgesellschaft“ eigentlich?

2. Der Begriff der Information als Problem

Dem Begriff Information begegnen wir historisch gesehen schon im Handel, Gerichtswesen und kirchlichen Leben des Mittelalters. Nach P. Toman ist der Ausdruck „Information“ zum erstenmal im Jahre 1274 in der Bedeutung eines Ensembles von Geschehnissen verzeichnet, die zum Beweis einer Straftat und zur Aufdeckung deren Täter führen.

Das Wort Information kommt aus dem lateinischen „informare“, das einführen in eine Form, Gestaltgebung, formieren, gestalten, darstellen, vorstellen, eine Vorstellung schaffen bedeutet. Das lateinische „informatio“ drückt dann den Begriff, die Vorstellung, den Umriss aus.

Information ist ein Begriff, dem jeder Mensch tagtäglich begegnet. Es ist eine Universalerscheinung, nicht nur eng verbunden mit dem Alltagsleben des

Menschen und der Gesellschaft, sondern auch mit der lebendigen Natur, sie pflegt auch mit der unbelebten Natur verbunden zu werden.

Drahomíra Geistová führt in ihrer 1979 an der Philosophischen Fakultät der Karlsuniversität Prag verteidigten Dissertation 311 Definitionen und Erörterungen des Begriffs Information an. Gegenwärtig verläuft eine tschechisch-spanische Forschungsarbeit zu diesem Begriff, die an die erwähnte Dissertation anknüpft.

Die Polysemie des Begriffs Information führt viele Fachleute zu dem Schluss, dass wir eigentlich nicht wissen, was Information ist, oder sie kommen sogar zu der Meinung, dass die Information überhaupt nicht existiert. Sie fragen sich : Ist der Begriff Information nicht den Begriffen Fluidum, Flogiston u.ä. ähnlich? Ist Information nicht eine Mystifikation?

In der Praxis und in der Wissenschaft wird jedoch dieser Begriff immer mehr benutzt, das Leben braucht ihn. Der Informationszutritt als Bestandteil des systemhaften Zutritts ist eines der immer mehr in der Wissenschaft benutzten Rationalisierungsinstrumente. Es entfalten sich solche Wissenschaftsdisziplinen wie die Informationstheorie, die Informatik und die Informationswissenschaft. Die Philosophie reiht den Begriff der Information zu solchen universellen Begriffskategorien wie Materie, Energie und Bewegung.

Gegenwärtig wird der Begriff der Information in vier Grundbedeutungen gebraucht als

1. Maß für den Grad der Beseitigung von Mißordnung (Entropie), Maß der Organisation im System,
2. menschliche Kenntnisse, Erfahrungen und Erlebnisse, als Zeichen in der Materie bzw. auf der Materie verzeichnet oder mittels Signalen und Impulsen in technischen Einrichtungen und im globalen Raum und im Weltall zirkulierend,
3. psychophysiologische Erscheinung und ein ebensolcher, im menschlichen Bewusstsein verlaufender Prozess und
4. Ausdruck der Verschiedenartigkeit in lebenden Objekten, eventuell auch in der unbelebten Materie.

Lassen wir die Information in der kybernetischen Auffassung als Maß der Beseitigung der Mißordnung (Entropie), als Maß der Organisation im System beiseite. Auf der Grundlage dieser Auffassung entstand wie bekannt Shannons Informationstheorie, die heute in der Welt ihren festen Platz innehat und eine reiche Forschungsbasis besitzt. Ihre Entfaltung ist die Triebkraft der schnellen Entwicklung der Computer- und Telekommunikationstechnik (computer science, Informatik).

Mit dieser ersten Auffassung hängt auch eng die zweite zusammen: Die mittels Zeichen, Signalen und Impulsen in der Materie und Energie verzeichnete Information. Die Möglichkeiten der Aufzeichnungen einiger Ergebnisse der menschlichen Kultur (hier weit gefasst) und deren Verbreitung begannen in der menschlichen Gesellschaft seit der Entdeckung der Schrift vor mehr als 5 000

Jahren resp. seit der Schaffung der ersten Höhlenzeichnungen vor vielen Tausenden Jahren aufzutreten. Heute erreichen diese früher ungeahnten Aufzeichnungsmöglichkeiten und deren Verbreitung mittels der Computer- und Telekommunikationstechnik ihren Höhepunkt. Es handelt sich um potentielle Informationen und Daten als eine Art der in technischen Einrichtungen zirkulierenden potentiellen Information. Deswegen können die so aufgezeichneten Informationen gemessen werden (Bits, Bytes), besonders in der zweiten Hälfte des 20. Jahrhunderts kann man deren riesiges Anwachsen verfolgen. Ich nehme an, dass der Ausdruck „Informationsgesellschaft“, aber auch die sehr frequentierten Ausdrücke „Informationsexplosion“, „Informationskrise“ usw. aus der Begeisterung und der Blendung durch die neuen Möglichkeiten, die in dieser Hinsicht die technische Entwicklung der modernen zivilisierten Gesellschaft mit sich brachte, entstanden sind.

Es besteht kein Zweifel daran, dass in den sog. entwickelten Ländern diese Entwicklung früher nicht geahnte Möglichkeiten des Individuums für den Zugang zu Kenntnissen, Erfahrungen und zur Aufzeichnung einer unendlichen Menge der von Menschen ausgedachten Begebenheiten mit sich brachte. Ebenfalls ist der Einfluss dieser potentiellen Informationen auf den Menschen und die ganze Gesellschaft unbestritten. Es ändert sich die demographische Struktur der Gesellschaft, es kommt zu ihrer ökonomischen Entfaltung, es ändert sich der Lebensstil der Gesellschaft und jedes Einzelnen.

Trotzdem bin ich überzeugt, dass für den Menschen die dritte Auffassung der Information als psychophysiologische Erscheinung und Prozess entscheidend ist.

Der Begriff der Information ist wertneutral. Den Wert verleiht der Information erst ihr konkreter Empfänger, je nach dem, was sein Bewusstsein empfängt und in eine Erkenntnis verarbeitet. Für die Persönlichkeitsentwicklung des Menschen ist dabei eine Fülle äußerer Anregungen in Form von Informationen resp. eine ausreichend entwickelte „Vorratskammer“ an inneren, aus seinem Endozept stammenden Anregungen entscheidend für sein Denken und das folgende Verhalten und Handeln. Das Verhältnis zwischen dem Denken auf der einen Seite und dem Verhalten und Handeln auf der anderen Seite ist allerdings nicht geradlinig, einfach und unproblematisch. Die Möglichkeit der Auswahl aus einer ausreichenden, aber erträglichen Menge potentieller Informationen ist für den Menschen wohlthuend, nicht aber entscheidend. Übertritt der Empfänger diese erträgliche Grenze, hört er auf, sich zu orientieren, verliert er die Fähigkeit der Auswahl, vermindert sich seine Fähigkeit, diese Information in Erkenntnis zu verarbeiten. Darüber, welche Information der Mensch aus der heftig anwachsenden Menge der potentiellen Informationen auswählt, entscheiden seine charakterlichen und moralischen Eigenschaften, besonders sein Wille zu erkennen und seine Fähigkeit, die aufgenommene Information zu werten (ihr einen Wert zuzuteilen) und diese in Kenntnisse, Erfahrungen und Empfindungen umzugestalten.

Die Persönlichkeit des Menschen wird also nicht primär durch die Informationsaufnahme herausgeprägt, sondern durch deren Verarbeitung und durch die Denkprozesse.

Darum bin ich der Meinung, dass für die weitere Entfaltung der menschlichen Kultur nicht die Informationen und die sich ständig vervollkommnenden technischen Mittel zur Informationsübertragung in der Gesellschaft entscheidend sind, sondern der menschliche Wille, die Moral und die Fähigkeit des Menschen, aus der anwachsenden Menge potentieller Informationen eine Auswahl treffen zu können. Darum kann ich auch unter diesem Aspekt behaupten, dass die kommende Gesellschaft eine andere Bezeichnung als die der „Informationsgesellschaft“ tragen sollte.

Über den Charakter der künftigen Entwicklung der menschlichen Gesellschaft wird mit unsichtbarer Hand das menschliche Denken und Handeln entscheiden.

3. Vom Verständnis des Menschen zum Verständnis der Natur

Es existiert allerdings noch ein Grund, warum ich die Bezeichnung der antretenden Gesellschaft als Informationsgesellschaft für ungeeignet halte. Dieser Grund betrifft die vierte, bisher wenig ausgearbeitete Auffassung der Information als Ausdruck der Verschiedenartigkeit und Kooperation in belebten Objekten resp. in der unbelebten Natur, deren unteilbarer Bestandteil wir waren, sind und immer sein werden.

Die Information als psychophysische Erscheinung und Prozess erschöpft sich nicht nur in der zwischenmenschlichen direkten und indirekten Kommunikation und in der Kommunikation des Menschen mit sich selbst. Es handelt sich um einen Informationsaustausch zwischen uns und der Biosphäre. Die (vereinfacht technische) Zivilisationskomponente der menschlichen Kultur mit ständig anwachsenden gegen die Natur gerichteten Elementen führt zur Vertiefung des Konflikts zweier sich entwickelnder Ordnungen, der Ordnung des Planeten Erde und der Ordnung der menschlichen Kultur. Dieser Konflikt äußert sich und reift im globalen Maßstab als Konflikt natürlicher und kultureller Evolution (J.Šmajs). Erstrangig bedroht er nicht die Natur, die Biosphäre, den Planeten Erde, den es hier Milliarden Jahre vor der Existenz des Menschen gab, sondern er bedroht ernsthaft die menschliche Kultur.

Die Informationskommunikation schafft die Voraussetzungen für das Verstehen und das Verständnis und zwar nicht nur der Menschen untereinander, sondern auch für das Verständnis der Menschen und der Natur (der Biosphäre). Es genügt nicht nur die konstitutiven Informationen (den genetischen Code) der menschlichen Kultur, sondern auch die konstitutiven Informationen (den genetischen Code) der Natur zu verstehen. Die Überwindung des oben angeführten Konflikts lässt sich in der Synergie beider Arten der Ordnungen, beider Arten der konstitutiven Informationen, erblicken.

4. Schlussbetrachtung

Der französische Schriftsteller André Malraux sagte, dass das 21. Jahrhundert spirituell oder überhaupt nicht werde. Anders ausgedrückt: Das 21. Jahrhundert wird entweder ein Jahrhundert der geistigen Wiedergeburt mit einer neuen Ethik menschlicher Beziehungen und Beziehungen des Menschen zur Biosphäre, wie es die amerikanischen Prognostiker J. Naisbitt und P. Aburdeen in ihrem Werk „Megatrends 2000“ vorausgesagt haben, oder es wird überhaupt nicht sein.

Literatur

CEJPEK, J. *Informace, komunikace a myšlení. Information, Kommunikation und Denken* (in Druck)

ŠMAJS, J. *Ohrožená kultura. Bedrohte Kultur. 2. Ausgabe*, Praha, 1995

Manpower Requirements in the Emerging Information Society of the Baltic States

Elena Macevičiūtė

Faculty of Communication, Vilnius University,
Universiteto 3, Vilnius, Lithuania,
e-mail: elena.maceviciute@kf.vu.lt

1. Introduction
2. Methodology design and execution
3. Demand for information specialists in the main sectors of employment
 - 3.1. Information service sector
 - 3.2. Information specialists in industry
 - 3.3. Library sector
4. Qualitative requirements for information specialists
5. Conclusion

Summary

The presented work is based mainly on the results of "Survey of Information Specialists Labour Market in the Baltic States", a research project which was undertaken by the Faculty of Communication at Vilnius University and the Department of Information Studies at the Tallinn University of Educational Sciences. The objective of the research was to establish the main tendencies in the labour market for the information specialists that will enable to plan the studies and number of students for the institutions of higher education. According to this objective three main sectors of employment are selected: a) emerging sector of information service; b) established sector of libraries; c) information management in different sectors of economics. The author introduces methods of research and main findings.

1. Introduction

The presented work is based mainly on the results of "Survey of Information Specialists Labour Market in the Baltic States", a research project which was undertaken by the Faculty of Communication at Vilnius University and the

Department of Information Studies at the Tallinn University of Educational Sciences according to the recommendations of the UNESCO/IFLA/EUCLID Workshop on Education and Training of Information Specialists in Eastern Europe and CEI (UNESCO 1995). The Workshop was supported by ASSISTANCE. The need of labour market in information specialists was never researched before in the Baltic States. The project was carried out in 1995-1997, additional data was collected in spring of 1998. The data from literature and periodicals about Latvia is used to complete the picture on the Baltic states.

The objective of the project is to establish the main tendencies in the labour market for the information specialists for a long term planning of higher education in library and information science in the Baltic states, as well as to find out and define new areas and possibilities for information specialists in the emerging information sector. The project includes three basic steps and involves:

- 1) the development of methodology for research of the changing labour market in information sector;
- 2) data collection and creation of a data base;
- 3) analyses and comparison of the data and recommendations for the educators.

2. Methodology design and execution

The very start of the work on the project was marked by considerable conceptual problems. The obvious and recently clearly defined profession of a librarian and information specialist has become an opaque concept without clear job description or professional obligation. The second difficulty was the establishment of the sectors of employment for the information specialists that should be covered by the project. The preliminary exploration of the job and profession classifications used by the labour exchange and other employment institutions in Lithuania and Estonia and study area classifications of the institutions of higher education or research have presented a very ambiguous picture. Even the librarian's profession may be found under different chapters in the same classification, not to speak of the different kinds and aspects of the information professionals. Only the Classification of Economic Activities (NACE) constantly places the information and library field in the "Other communal, social and personal services" area. To define the attitudes of the participants of the project towards these main problems was essential for the further design of the methods and data collection.

The current need to identify an information specialist or a new information professional is justified by emergence of information intensive economies throughout the world. The concept of the profession changes according to the development of the information sector and under the influence of the changing needs of society as well as requirements of the new information technology. According to some concepts the information sector is treated as the largest

sector in economics and most of the workers as doing one or another kind of information job. However, this is not entirely correct and cannot be used as a basis for our research.

The broad concept of information specialist requires a major survey of demand and professional requirements in a huge sector of economics and is unrealistic with regard to available resources. The objective of the present research is to establish the main tendencies in the labour market for the information specialists that will enable to plan the studies and number of students for the institutions of higher education. According to this objective three main sectors of employment are selected:

- a) emerging sector of information service (governmental and private);
- b) established sector of libraries;
- c) information management in different sectors of economics.

Each sector presents special problems and difficulties for researchers and requires a slightly different approach though the main methodology should remain the same.

The objective of the first stage was to work out the complex of methods of research that is suitable for the establishment of the main trends in the labour market for information professionals in the Baltic states. It could be repeated and applied in other Central and European states and give results compatible with the results in other parts of the world.

The greatest challenge is the changing environment and society of the newly emerged democratic states and the rapid growth of information sector. The change from the centralised economics to decentralised systems and the liberalisation as well as the growth of the private sector change the information activities and services very quickly. It is impossible to predict the development of the needs and requirements for modern jobs in information field for certain as the investments into information technology of the country and from abroad is constantly increasing.

The main activities carried out with regard of the methodology development was the analysis of the similar research carried out in other countries, review of the available information resources and the preliminary examination of the situation in information sector.

The main conclusion reached by the working group is as follows: the diversity and quick changes of the information sector require constant monitoring and data collection about the quantitative demands and qualitative requirements for information specialists using a variety of research methods. Only this way can produce reliable and usable data for forecasting of the demand, planning of professional education and curriculum development.

For the current research the following methods were used:

1. Analysis of the statistical sources and documents used in labour exchanged systems. The main result of this activity was to define the status of the information specialists among other workers and the reflection of the job demand for these specialists in the labour exchange systems of the CEE countries.
2. The analysis of the advertisements in the periodical press. The result showed awareness and general job definitions of the information-related jobs as reflected through mass media (indirectly showing the status of the information specialists in society).
3. The analysis of the growth in information services, businesses and systems.
4. The survey of the employers (questionnaire) was the main method for collecting data about the possible changes in the market and the present requirements for information professionals. It is implied that this way the data about the people occupying the jobs of information professionals at present was also collected.

The described methods were intended to complement each other and were used selectively for different parts of information sector research.

The questionnaire suitable for the collection of the data in Estonia and Lithuania and applicable to the different parts of the surveyed manpower parts in the information sector was prepared on the basis of UNESCO guidelines (UNESCO 1986). The questionnaire was prepared in two parts (and third part was added for the survey in library sector):

- PART I consisted of the questions the answers to which supply data for the qualitative evaluation of the labour market for information specialists;
- PART II was designed to supply qualitative data on the skills and knowledge demanded by the market from information specialists. The list of competencies was modified from the research made in Latin America (Paez-Urdaneta I. 1992).
- III PART is designed only for the survey of librarians and serves to gather data about the attitudes and orientation of present professional librarian in the context and main problems of the library work as well as the attitude towards the professional development and activity.

The second objective was to collect reliable data that could be used to characterise the labour market for information professionals in Lithuania and Estonia. The questionnaires were disseminated among information services providers in April. One third of them was returned and processed. The first data from the library sector was collected in October. The collection of data from the third sector (information management in other economic sectors) was finished in May, 1997.

3. Demand for information specialists in the main sectors of employment

The labour market in the Baltic states is a comparatively new officially recognised phenomenon. The organisation of information flows in the labour

market helps employers to fill vacancies in time and the employees to find jobs. An effective system of employment should be based on the active processes of job seeking and information support for these processes. However, the information factor is not considered to be important in the regional employment programmes. The labour exchanges in Lithuania, Latvia and Estonia are the main mediators between employers and employees from 1991. These institutions were established according to the state of economics and to the experience of countries with developed market relations (Sweden, Denmark, Germany). It is the main institution gathering information about labour demand and supply. The data about employment are received in the surveys, questioning the citizens. The surveys are carried out in according to the recommendations of the International Labour Organisation (April 1994; February and May 1995; June 1996). The main trends in the labour market of the Baltic states are: the growth in unemployment, decrease in vacant jobs, increase in regional differences in unemployment, the growth of the number of unqualified young people, and increase in the numbers of registered unemployed persons. The main growth of jobs takes place in the private sector, where specialists knowing foreign languages and working with computers are in great demand.

According to the households survey in 1995 the main channels of getting information about jobs are:

- friends and acquaintances (60%);
- labour exchange (50%);
- advertisements in newspapers (40%);
- direct contact with employers (20%);
- organisation of individual business (2.3%).

Employers looking for employees use:

- labour exchange data (50%);
- recommendations of the enterprise employees (36.4%);
- lists of previous employees (27.6%);
- advertisements in newspapers (13%).

According to the results of the surveys we used the labour exchange data and the advertisements in newspapers to define the information specialists demand and supply situation. The Labour Exchange in the Baltic states do not separate information specialists as such. Information work does not figure in the Classification of Economic Activities either, most probably because of its novelty. The data bases on employment mark out Information activity as a part of the sections "Computers and related activity", library work is a part of the section "Recreation, entertainment, cultural and sports activities" not to speak of other sections in which information specialists are disguised as well. In fact the Labour Exchange information system does not allow to evaluate the demand and supply situation for information specialists even according to institutional criteria.

The institutional criteria (the advertisements of libraries or librarians, information service firms and centres, enterprises looking for specialists for information

departments) were used to evaluate general labour demand and supply situation as reflected in the Lithuanian periodical press. Three main dailies, two regional newspapers and three advertising newspapers were selected. The main criteria for the selection of the papers was the existence of a special column for job advertisements (job seeking as well as advertisements about vacancies). The advertisements published from 1995/01 to 1996/02 were retrieved. The repeated advertisements were treated as separate ones as they reflected an unsatisfied demand. It is clear that information specialists do not generally advertise their availability for work in the newspapers: only one such advertisement from a librarian was found in a regional newspaper (*Sekunde*).

The demand for specialists in information sector was far greater. All in all 98 advertisements were found in surveyed newspapers. The largest number of advertisements offering jobs for information specialists was found in January, 1996 (25). A general trend of slight growth towards the end of the year 1995 and at the beginning of 1996 emerged in all three types of papers (1-2 ads per month). However, in some papers the number of advertisements for information specialists is very low, or they do not appear at all. The general overview of information specialists' labour market shows that the awareness of existence of this special type of professionals and the need for them is growing in Lithuania. However, there is an evident lack of general information sources about this specific section of labour market.

3.1. Information service sector

The direct indicator in the information service sector growth and the indirect indicator of the labour market for information specialists situation is the number of information service companies and centres.

One of the features that characterises the growing Baltic market of information services after 1990 is differentiation of services according to changing demands of the users. Before 1990 most information centres offered to their users scientific technical information and served the needs of research and innovation activities. At present the market demands mostly services for growing Lithuanian and Estonian businesses. Bigger enterprises and companies can afford to establish a department of information usually combined with public relations activities or at least to hire an information specialist, however, a growing number of small and medium businesses turn to specialised agencies. The need was quickly perceived and the number of institutions and companies that offer information services started growing rapidly. Some of the previous ones have converted their services and turned to serve the most prospective users – businessmen (Lithuanian Institute of Information, Lithuanian Technical Library).

The situation in the information services sector is rather complex and difficult to sum up, because of the indefinite usage of the term "information services" and the wide range of the companies offering them. For example, in Estonia in 1997

there were 156 firms according to the REG-2 (electronic database of Estonian firms) offering business consultations. Other companies are selling information technology and provide services connected with the exploitation of those. However, some of them are broadening the scope of their activities in connection with Internet (Virkus S., Maceviciute E. 1998). We have tried to establish the rate of growth of the business information companies in Lithuania from 1992 to 1996 checking the data provided in companies' registers and catalogues. Those that provide to the customers collection, processing and dissemination of economic, industrial, commercial data, market research, etc. are considered as the providers of information services. The data shows that the number of these companies have grown twice in Vilnius from 1993 (from 24 to 58) and the growth rate is rather steady every year (from 8 to 15 per year). During 1996 that number increased to 92 (in January, 1997).

The survey of employers was conducted at the beginning of 1996 and reflects the labour market situation in 1995. The questionnaires were sent to 58 companies and institutions. 30 of them returned the questionnaires after several reminding (51,7%). However only 28 were considered fit for the analyses (48,2%). According to the returned questionnaires 36% of information service companies are creating data bases and disseminating or publishing information from them; 21% – creating data bases and providing information from them as requested; 21% – establish reliability of business partners; 21% indicated such services as marketing, information retrieval, publishing newsletters or information leaflets, etc.

The labour market for information specialists quantitative situation in this sector is represented in the summary *Table 1*.

Table 1. Number of employees, number of vacancies

	Types of services				Total
	DB, publish.	DB, retriev.	Business reliability	Others	
1. Staff employed	240	72	40	278	630
2. Vacant posts	10	4	28	8	50
3. Total posts	250	76	68	286	680
4. Vacancy rate	4%	5.3%	41.2%	2.8%	7.4%

The results indicate that there is a balance between specialists demand and supply. The percentage of vacancies in this sector is 7.4%. The least satisfied are the companies performing deep information analyses services in business sphere, i.e. demanding the specialists of the highest qualification. Total turnover in the information service sector is approximately 19% and in some sections reaching up to 35%. This turnover is rather high. Further analyses of the data allows to state that the large arrivals rate is equalled by the high departures rate in the information services sector, especially in business information services. One can indicate that most probably these services lack

good professionals who are qualified for the job. The financial considerations of employees in these services could not serve as a reasonable explanation for quitting the job. The new arrivals are most often attracted by high salaries and expectations.

The recruitment need in five years is quite obvious from the data presented in Table 2 and indicates that the job market in this sector is unsaturated.

Table 2. Supply and demand in 5 years

	Types of services				Total
	DB, publish.	DB, retriev.	Business reliability	Others	
1. Present supply	240	72	40	278	630
2. Minus departures	300	70	70	160	600
3. Plus arrivals	350	110	100	10	570
4. Supply in 5 years	290	112	70	128	600
5. Present posts	250	76	68	286	680
6. Plus net growth in 5 years	62	54	70	40	226
7. Demand in 5 years	312	130	138	328	906
8. Demand in 5 years	312	130	138	326	906
9. Minus supply in 5 years	290	112	70	128	600
10. Recruitment need in 5 years	22	18	68	198	306

The answers to the questionnaire show that, at present, most often the information services employ specialists who have no professional education in information science: economists, journalists, engineers, teachers, designers, physicists, mathematicians, etc. The fact that this niche is filled by the job-seekers from different fields shows that the demand for information specialists is not satisfied.

3.2. Information specialists in industry

Despite the difficulties of the transitional period Lithuania remains an industrial state. Though the share of industry in the General Average Product has become smaller, it is 29% and is growing constantly. The process of privatisation according to the law has reached all the industries and regions. Main factors that influence the labour market in industry are: economic situation, structural and functional changes in economy and management, formation of market economy relations. The revival of the economy is manifesting itself in revival of the services that were closed down or squeezed during the first hard years of restructuring. Information services are among them. At present the need and significance of information in industry is rising.

However, no one has tried to survey the possibilities of this sector from the point of view of the labour market for information specialists. A first modest attempt was made in order to complete the picture of the whole situation at the end of this survey. From the surveys of the industrial and services sector in Estonia conducted in 1995 it was clear that the greatest part of small and average-sized businesses do not employ information specialists permanently. Mainly they use the possibilities offered by information services and centres. However, the larger and stronger industrial enterprises establish either posts or departments to fulfil the information needs of the organisation. That was the main criterion for selection of the respondents for the survey. In Lithuania the biggest and best enterprises are announced in the main business weekly "Verslo žinios". The list of 200 best and largest enterprises in 1996 was published in February, 1997. 90 (45%) enterprises were randomly selected from the list and the main data on them was gathered from the registers and catalogues. 40 (44%) of the surveyed enterprises from different towns have returned the questionnaires by the end of April, 1997.

The analyses of the data has showed that practically there are no vacancies for information specialists – the posts of that kind are filled. The negative net growth of posts (diminishing) in three years may be explained by the fact that the climate in industry is unstable and too many factors effect the enterprises. The employers do not intend to create new posts in administration or services. They are too concerned with general survival. However, one has to have in mind that this assumption may be wrong. It is possible that the present level of information services in the surveyed enterprises is satisfactory, or that the employers do not see the necessity of creating this kind of new posts. The conducted survey does not allow to give the final conclusion on the basis of available data.

The arrivals rate among information specialists exceeds the departure rate. Some of the enterprises have created totally new posts for this kind of employees. The overall results show that the industry as a labour market for information specialists is not growing at present and the recruitment need is rather low. It may be satisfied without additional educational effort. However, the survey of this sector has revealed its difference from the other two sectors. The employers in it have not met with the concept of the information specialists and are not used to it. Besides, they are not primarily interested at the problems of recruitment of these specialists and are not ready to spend time thinking about this. This was not taken into account before the start of the survey. For further research in this sector it would be worth looking for a different method of data gathering or at least to adapt the questionnaire to its peculiarities. However, the evaluations of the main competencies of information specialists in this sector are very useful for further recommendations:

3.3. Library sector

According to the international standard ISO 2789: 1991 (E) prepared after the UNESCO guidelines "a librarian" is a person who works in a library, receives a salary and is educated professionally to do this work (on the job or at the educational institution).

Labour market of librarians in Estonia and Lithuania at present is a problematic object of research for several reasons: 1) the relevant statistical data necessary for deep analyses is scarce or non-existent; 2) libraries are still undergoing major changes; 3) the system of library jobs may fall under direct change as well. It is difficult to predict which system of job categorisation will eventually be implemented. Statistical offices and Ministries of Culture provide data on different types of libraries. However, staff data is represented poorly and only with regard to educational qualifications. It allows to make only relatively simple and routine forecasts.

There are 3135 librarians working in Estonian libraries. 48% of all librarians are professionally educated, more than half of them have university degrees. With respect to different types of libraries the picture is somewhat different. Of 1184 librarians at work in public libraries 320 have university degrees and 446 have college degrees. So six out of ten librarians have professional education. The education qualification of librarians in rural areas are considerably lower than in urban areas.

Table 3. Standards for education in libraries urban and rural areas

		Urban (%)	Rural (%)
Professional education	university degree	37	14
	college degree	3	2
	secondary education	33	39
Education in other fields	university degree	13	8
	secondary education	14	37
	Total	100	100

Of librarians in research and special libraries 38% have professional education and in school libraries 37% of librarians have professional education (a 5% rise compared to previous years). Although the number of librarians with higher professional education has been constantly increasing in recent years, the librarians without qualification still make up more than a quarter of all librarians. The bigger the library, the higher is the number of employees with university and college degrees in other fields (computer specialists, economists, sociologists, lawyers, psychologists, etc.) (Papp U., Markus T. 1996). In Lithuania the analyses of the staff in the public libraries shows that the number of the people working in the libraries is rather stable and the number of posts is growing slightly. The difference may be explained by the fact that the salaries in the libraries are low and the same person tends to occupy more than one post.

Table 4. Library posts and staff in Lithuania

1994		1995		1996	
Pers.	Posts	Pers.	Posts	Pers.	Posts
2659	3314	2665	3317	2664	3361

84.4% of library workers in the libraries of the Ministry of Culture system have professional education (19.48% – university degree, 52.36% – vocational library education). The situation is different looking at the school libraries. In 1996 they had 1775 staff members, 339 (19.09%) of them with a university degree and 32 (1.8%) with a vocational education.

The questionnaires were distributed among 114 libraries and 60 (52.63%) of them have returned the questionnaire. Among those that returned the questionnaire 7% – special libraries; 14% – university libraries and 79% – universal public libraries. No questionnaires were distributed among school libraries. The answers have shown that 18% of libraries have created at least one and 20% plan to create additional posts (from 1 to 5) in next three years, 17% have deleted posts in 1996, but no one plans to delete them in the future. The departure of librarians was quite significant in big cities, yet the new employees were not difficult to find there. The libraries in smaller towns mostly experience loss for natural causes (retirement, death). Most libraries claim difficulties in finding young specialists or specialists with higher education (63%). The age of the librarians may be envisaged as the factor that will cause greatest problems in the future. 22% of the staff members are over 50 years, 7.05% among them over 60 years old. 49.19% are the people in between 36-50 years. Having in mind that the majority of library workers are women (retirement age – 60), in ten years every fifth librarian will retire and should be replaced. There is no possibility to make a more precise forecast of the labour market for the librarians at present if the library posts system will undergo major change.

4. Qualitative requirements for information specialists

The quality requirements for modern information specialists may be determined from the answers to the questionnaire (part II). The most valuable competencies differ slightly in the surveyed sectors. The rating interval was a scale from 0 to 3 (3 being the highest rating for the competence).

Table 5. What competencies are most important for your institution?

Competencies	Rating in		
	Information service sector	Library sector	Industrial sector
1. computerised information storage and retrieval	2,9	2,0	2,8
2. systematic search of information	2,7	2,5	2,8
3. information resources management	2,6	2,3	2,3
4. selective dissemination of information	2,6	1,7	2,1
5. promotion of organisation's info-resources	2,5	1,8	1,8
6. use of external databases	2,3	1,8	2,5
7. production of commercial data bases	2,3	0,8	2,1
8. use and design of information systems	2,1	1,5	2,0
9. design and marketing of infoproducts	2,1	1,4	2,8
10. forecasting and current planning	2,0	2,2	2,5
11. in-house data base design and development	2,0	1,6	2,5
12. training in inforesources and technologies	2,0	1,8	2,1
13. reviewing of specialised literature	1,9	1,6	1,8
14. optimisation of information flows and functions	1,8	1,2	2,3
15. production of abstracts and reviews	1,7	1,7	2,0
16. preparation of reports and newsletters	1,7	1,8	2,1
17. creating bibliographies, indexes and catalogues	1,5	2,4	1,2
18. infosupport to decision making	1,5	1,4	2,3
19. other (indicate)	-	local history work; event organisation	-

The differences are defined by the nature and objectives of the information specialists working in different sectors. Libraries are non-commercial institutions, the usage of the modern information technology is at the lower level, and the functions as well as the work processes in libraries have special features. Therefore creation of bibliographies and catalogues in the libraries has much higher rating. The high rating of optimisation of information flows and functions and infosupport to decision making in industrial sector is defined by special features of information work in these organisations similar to selective dissemination of information and promotion of organisation's inforesources in information service sector. However, there are competencies: systematic search of information; information resources management; and computerised information storage and retrieval - that are highly rated in all sectors.

The attributes in information professionals are slightly different in information service and library sectors. The data received from the industrial sector in answer to this question is incomplete and is not presented here.

Table 6. What attributes do you prefer in information professionals?

Attribute	Rating	
	Information service sector	Library sector
verbal and written communication skills	2,7	2,5
need of self-improvement	2,6	2,6
English language	2,5	2,8
friendliness	2,5	2,1
ability to use the computer	2,4	2,8
foreign language	2,3	2,8
management skills	2,1	2,3
inclination to science and technology	2,1	1,5
higher education	2	2,5
knowledge of economics	2	1,5
education in a second area	1,9	2
knowledge of managerial concepts	1,8	2,4

Table 6 shows that though the computer usage at present is not as high as in information service sector, the ability to use the computer is considered to be very important. Only the foreign language skills are valued at the same rate. Need for self improvement is as ranked higher than university degree and communication skills.

The third part of the questionnaire shows the attitudes of the present library staff towards the context and main orientation of the libraries in society. This survey covered 80 respondents in Estonia and 67 respondents in Lithuania. 32.5% of them were working in academic libraries, 11.3% in school and children libraries, 56.3% in public libraries. Half of the respondents are library leaders. In Lithuania this part of the questionnaire was distributed during the annual conference of the Lithuanian Librarian's Association. The indicators of professional attitudes may be categorised into three groups:

- relating to internal factors of library
- relating to external factors
- relating to activities of librarians themselves.

Both in Lithuania and Estonia more than half of respondents give high priority to all presented statements. Appreciation of the role of library by state and local authorities received highest ratings (83% of respondents gave the highest ranking 5 to it). In Lithuania the next highly ranked statement concerned the need for a librarian to know the new ideas about library development and experience (70%). The existence of the strategic plan and understanding social environment was not important for 25% of respondents in Estonia, however, in Lithuania only 6% of respondents indicated that strategic plan is not important, 6% had no opinion if it is necessary to design active and purposeful

development of libraries and 30% had no opinion if librarian should understand social environment.

Evaluating present and future problems the librarians have not shown great differences between present and future problems. Most of the problems were derived from the lack of funds, including low competence and motivation of librarians, lack of foreign language skills and young specialists, poor acquisition, low automation and equipment level. However, conservative attitudes of librarians, indifference, lack of normative documentation and standards, lack of co-operation among libraries, service quality and variety were mentioned as well. Among the Estonian respondents 40% were unable to define the critical points in their library. All Lithuanian respondents indicated the problems in their libraries and common problems for the libraries in Lithuania. The difference may be accounted to the fact that Lithuanian survey included the most active part of the professional body.

5. Conclusion

The answers to the questionnaire show that, at present, most often the information services employ specialists who have no professional education in information science: economists, journalists, engineers, teachers, designers, physicists, mathematicians, etc. Information service and library sectors form a constant and growing market for the labour of information specialists. For the lack of information specialists this niche is filled by the job-seekers from different fields. The other potential sectors should be monitored and closely surveyed using special methods. The most important competencies (as measured on a three-point scale) are those from the information specialist's repertoire. Computerised information storage and retrieval as well as traditional information search abilities top the list of information specialist's competencies. Ability to use the computer is highly valued as an attribute of an employee though it does not outrank such qualities as oral and written communication skills, need for self-improvement, friendliness and knowledge of English language. The attitudes analyses has showed that professional skills and knowledge as well as the ability and willingness to acquire knowledge permanently is given a high priority among professional librarians.

According to the results of the research Vilnius University has created a new Bachelor's degree programme in Business Information Management as the need for a specialised expertise in this field was indicated by the high turnover rates revealed by the research in the information services sector. The ranking of the main competencies and skills in different sectors of the labour market were taken into account in changing curricula (e.g., introducing a Communication Skills module) for the information specialists and librarians not only at the Faculty of Communication of Vilnius University and Tallinn Pedagogical University but at the other higher educational institutions as well (e.g., in design and establishment meetings and curriculum development

activities for the International Centre of Information Management Services and Systems at the Torun University).

References:

[Paez-Urdanet I. 1992]

Paez-Urdanet I.: *To experience a connection: in search of a new information professional for Latin America*. In: *State of the modern information professional 1992-1993*. The Hague: FID, 1992, p. 33-35.

[Papp U., Markus T. 1997]

Papp U., Markus T.: *Is the library a learning organisation?* Tallinn, 1997, p.6. Manuscript.

[UNESCO 1986]

UNESCO: *Guidelines for conducting information manpower surveys: V. 1-2*. Paris, 1986.

[UNESCO 1995]

UNESCO: *Workshop on Education and Training of Information Specialists in Eastern Europe and CIS Countries, Bratislava, Slovakia, 14-18 November, 1994*. Final report. Paris, 1995, p. 14.

[Virkus S., Maceviciute E. 1998]

Virkus S., Maceviciute E.: *Usage of Business information in Estonia and Lithuania*. In: *Information management in small and medium-sized enterprises*. The Hague: FID, 1998, p. 28.

Information Sector as Economical Category

Rudolf Vlasak - Stanislav Kalkus

Institute of Information Studies and Librarianship, Faculty of Philosophy,
Charles University of Prague, Czech Republic, Celetná 20, 116 36 Prague 1
e-mail: rudolf.vlasak@ff.cuni.cz, stanley.kalkus@ff.cuni.cz

INTRODUCTION

About at the break of sixties and seventies of this century the first functional and publicly used information system developed from originally experimental and non commercial applications of computer technology in storing and retrieval of document records. At the beginning these were mostly parallel editions of bibliographies in both printed and magnetic format. Thus developed on one hand data bases of national bibliographies derived from the MARC project of the Library of Congress and on the other hand special and specialized data bases such as CHEMBASE, INSPEC, ERIC or WPI. At the same time important commercial online information retrieval service centers, DIALOG, ESA-IRS, DIMDI, STN and others were beginning to operate. Also at that time thoughts about delineating new branch in world economy "the information sector" were springing up.

Information services then became relatively fast an economical sector. Investments into computers, telecommunications, and also personnel grew rapidly in organizations which undertook business in the field of information services. These then ranked between the most successful enterprises according to statistics reporting the annual growth of business turnover, and also of financial gains. These were on one hand the producers of data bases and on the other hand vendors of online retrieval services. This enterprising is at all times characterized by globalization of markets, and also by investments into research and development because the innovation cycle is exceptionally short and the competition great. All mentioned signs of this type of enterprise caused the creation of what is today described as INFORMATION INDUSTRY. We do not mean by that an industrial production of information technology but rather the industrial character of providing information services.

INFORMATION INDUSTRY AND MACRO ECONOMICAL CONCEPT

The foundation of information industry is traditionally formed by systems of services in the field of scientific and technological development and its economics. These are services which have developed from traditional library services, active in both science and culture. Today's complex conception of information industry adds first of all category of generated and retroactively effective in the process of economical management, social management, further so called news information, and information of daily need in public and private life of individuals.

As any other industrial field, so does the information industry has its own economy. After economy of "factory smokestacks" and mass, assembly belt production, comes according to Alvin Toffler so called symbolic economy. Author follows the thesis that such a great switch of most new opportunities from the sector of material production into the sector of creation of information, and this process taking place not only in rich industrial countries but even in semi developed parts of the world such as we witness now, does not require detailed statistical studies in order to realize that society is changing from industrial to post industrial. The terms "post industrial society" and "information industry" are usually used in connection with the problems of "information society".

The dependence of today's world economy on information and telecommunication technology (ICT) evokes ever increasing dependence and interconnection of universality in economic and information sector. These which gradually merge into one "information system". In order to understand this phenomenon we even may use Vickery's already classical definition of an organization of people, materials and machines serving the transmission of information from one person to other.

While philosophers and sociologists talk today in connection with application of new technologies and of organization of knowledge, about post industrial, respectively information society, Peter Drucker for example promoted at the end of eighties the term "post business society" to describe the coming level of civilization in developed countries, and was comparing like Toffler, the importance of information wave with industrial revolution.

Democratization in access to information naturally leads toward democratization of rules and regulations in enterprises. Already in the second half of the seventies a discipline called information economics was developing in United States. In their pioneer work some economists, such as Marc Porat in his work "Information Economy: Definition and Measurement" begun to question not only to what degree information profession participates in the total professional spectrum of the labor active society, but also begins to define types and kinds of employment positions of which the whole information field is composed of. Besides the agricultural, industrial, and the service sector, the information sector was also defined. From the approach and definition, whether of the categories of work performance, or professions and kinds of employment which belong into this sector, and which economists selected as one of less differentiated from the macro economical

research of main streams of 20th century development, it is clear that it is impossible to clearly define boundaries between activities which we include in this sector and other social activities. Nowhere it is specified what kind of information is included in these thoughts. Actually it is not even possible. The main goal of the social economic studies in this direction is following the developmental trends of society as a whole and from the resulting tendencies (not from exact statistical numbers) derive strategic economic reflections. Therefore the individual authors, although they differ in various aspects of definition of the contents of information sector, they basically arrive to same conclusions. Mostly they deal with macro structures and so called big numbers, while they do not differentiate the important specifics of following trends of information activities from other viewpoints such as individual information services, individual technologies, different users needs to information systems access.

Let us try to outline the information sector by wider information fields which it is possible to differentiate either from the view of reasons and goals of information activities, or from the view of their institutional base.

INFORMATION CIRCUITS OF THE INFORMATION SECTOR

The term "information circuits" must be regarded as a proposal of a certain convention. From the first half of the seventies it enters the subconscious not only of the expert public involved in different aspects of information activities, but also of the managers on different levels. Often is the information sector identified with only one of its components. From the viewpoint of econo-political organization and management prevails the understanding which stresses the importance of mass media, the problems of public informations, and the question related to mass media and related activities. From the point of view of development of science and technology and generally of the whole tertiary sphere, the field of scientific, technical and economical information, including library activities prevails in the understanding of the public information sector. From the management point of view this sector includes together with all aforementioned activities also information operatively functioning in management of all kind of business enterprises and activities.

For the purpose of researching the coherence of the development of information industry the entire information sphere can be divided into six basic sections. The main principle of this division, as it is applied in the systems of information and library education in masters and doctoral studies at the Institute of Information Studies and Librarianship, Faculty of Philosophy, Charles University is a functional division of individually different but often closely related information sections which can be identified in the present developmental stage of the whole information sector. In this manner it is possible to define the following six sections: mass media, archives, public management, economic information, scientific and technical development, and cultural information circle.

1. Mass media section

This section of the information sector is pragmatically aimed at providing information in its most general sense and that without regard to professional to professional background of users (in this case more accurately consumers) of information, and their social standing. From the institutional point of view all news agencies belong into this group: the daily press and popular magazines, radio, television, newsreel, However also information services dedicated to public use such as video text, its present formats on Internet and its multimedia environment of www belong here.

To make public - is the main reason for this activity and for outlay of funds to support information services in this section. Organization, methods, technology and the processing techniques and dissemination of information are part of this. And of course also the formats by which the individual information systems are presented are to be considered within this section. It concerns formats of maximum understandable expressions and information dissemination. In some cases artistic formatting of contents is used to be transmitted by information channels in order to influence the recipients of these information.

2. Archival collections information section.

In contrast to the first section it is possible to specify this section from a number of viewpoints - final delineated field of methods and means effective on the average in a whole scale of social activities. Archives organize the information collections first of all to protect the contents and only secondary they allow selected access according to special information needs. Part of this section understandably belongs to non-public information services, however substantial part of archive and documentation can be included in the public information sector.

Main mission and reason for existence of these information activities can be simply expressed in one word: to preserve. All methodology and working techniques with information in this section is subordinated first of all to preservation and safeguarding of information documents, and also to the fact that any of the documents are retrievable and accessible at any time regardless of the time element connected with the origin of the information.

3. Public management information section

Considering the developments in the major part of the world toward democratic government and increasing participation of public in the management of state and also regional affairs, this section concerns more and more information of public type. Furthermore this information section grows in the last year into the field of administrative information.

Main aim of information activities and reasons for creation and maintenance of different systems and services within this information section is to secure qualified means of management of complex social economic and political systems, functioning of course on democratic principles in the environment of economical market relations.

4. Economic information section

The nucleus of information activities in this section is to support enterprising. Both from the angle of securing information for enterprise subjects in their system surroundings (that is supply and market areas), and also from the angle of the consumers providing information on goods which the enterprise sphere offers. Typical information activity which represents ever growing part in this information section is the management information system. Traditional fields of this section are stock market information, company and so called office information systems taking a maximum advantage of online information services, and paperless business and banking communication.

Here belong the macro economical indicators just as well as analytical indicators of company management. Furthermore there is a growing number of data bases containing detailed data on economical subjects, beginning with their names, addresses, names of company officers to a detailed listing of products and their parameters, or services which they provide.

5. Science and technology information section.

The term "scientific and technical information" (although sometimes named differently) is understood worldwide as complex of traditional library and bibliographic activities together with additional research and analytical activities serving primarily all creative activities, such as research, development, construction (design), education, health care, but also strategic management, prognosis, and planning. Information methods derived from services of special libraries, research libraries, academic libraries, information and documentation centers pass through all special activities and serve to elevate specialized levels of methodology and technology of providing information services in all other information sections. Very specific sign of this section is a strong orientation toward documents as a main source of information. Their collection, storage, and access by way of loans does not have analogy in so far characterized information sections.

The main goal of the processing activities and services of this information section is to support development of science and technology. The outstanding characteristics of this of this section is its mostly international level, just as international in character is the science and technology.

6. Cultural and art information section.

Without doubt activities which secure safekeeping and access to artistic works of all kinds and to other cultural properties belong to this information section. Presently the sphere of traditional public libraries with predominance of artistic, cultural, and general educational collections in form of printed materials is being widely unfolded. Thanks to Internet, systems are being developed which make encyclopedias, picture catalogs and reproductions of work of art accessible through the computer networks. Computer technology is the base of videotheks, phonoteks, and it takes its place generally as a foundation of information methods for accessibility of collections of these libraries. Besides computer media offering culture and art, audio recordings, video recordings, and optical discs are at home at the libraries for some time. These are made accessible to users of institutions in same manner as books and periodicals. Accessibility of culture to wide public in full meaning of this term is also the basic factor of this information section.

CONCLUSION

The developmental tendencies of society today are generally connected with information activities. We believe that it serves a purpose to categorize these activities and by that increase possibilities of an analytical view of resolving problems of present development of information sector and its increasing importance in economics and politics of this world. In this spectrum of opinion it is possible to underline those fields of information activities which from the viewpoint of researching actual connections of technological development and democratization of access to information are acceptable for the outlined views of these sections. The outline of the section boundaries may be only a framework without defining the definite borders but rather watch for their overlaps and mutual relations.

BIBLIOGRAPHY

1. Toffler, Alvin. *The third wave*. (New York: William Morrow, 1980)
2. Vickery, B. C. *Information systems* (London: Butterworth, 1973)
3. Drucker, Peter. *The New Realities: in Government and Politics, in Economics and Business*. (New York: Harper and Row, 1989)
4. Porat, Marc U. *Information Economy: Definition and Measurement*. (Washington, DC: U.S. Dept. of Commerce, 1977)

5. Vlasak, Rudolf. *Informacni Prumysl a Informacni Spolecnost (Information Industry and Information Society) in Acta Economica Pragensia*, No. 1, pp 147-177 (Praha, 1977)

Schutz von Daten und Schutz von Wissen in der Informationsgesellschaft

Gerhard Reichmann

Karl-Franzens-Universität Graz
Institut für Informationswissenschaft
Universitätsstraße 15 / F 3
A-8010 Graz

email: gerhard.reichmann@kfunigraz.ac.at

Inhalt

- 1 Einleitung
- 2 Schutz von Daten
 - 2.1 Rechtslage in Österreich
 - 2.2 Rechtsvergleich zwischen Österreich und Deutschland
 - 2.3 Probleme und Anpassungsbedarf
- 3 Schutz von Wissen
 - 3.1 Rechtslage in Österreich
 - 3.2 Rechtsvergleich zwischen Österreich und Deutschland
 - 3.3 Probleme und Anpassungsbedarf
- 4 Resümee

Zusammenfassung

Der Schutz personenbezogener Daten vor Mißbrauch hat ebenso wie der Schutz geistiger Schöpfungen - also von Werken, die durch Wissen entstehen - eine technische, eine organisatorische und eine rechtliche Komponente. Im vorliegenden Bericht wird die rechtliche Komponente dieses Schutzes behandelt. Die einschlägigen Rechtsvorschriften dazu finden sich einerseits im Datenschutzrecht und andererseits im Urheberrecht. Zunächst werden die derzeitigen Eckpfeiler dieser zwei Rechtsbereiche nach österreichischem Recht vorgestellt. Anschließend folgt eine Gegenüberstellung der entsprechenden deutschen Rechtsvorschriften, um in der Folge aktuelle Probleme und notwendigen Anpassungsbedarf innerhalb dieser zwei Rechtsmaterien, der sich vor allem aufgrund des Wandels der Gesellschaft von einer Dienstleistungs- zu einer Informationsgesellschaft ergeben hat, sowohl für Österreich als auch für Deutschland zu diskutieren.

Abstract

This paper deals with the legal aspects of the protection of the privacy of personal data and of the protection of mental creations. At first the essential parts of the Austrian data protection law and of the Austrian copyright are presented, followed by a comparison with the corresponding German laws. Then current problems of these two fields are discussed. Many problems are a result of the present change from a service society to an information society, where information is the most important good.

1 Einleitung

Vergleicht man unabhängig von länderspezifischen Gegebenheiten das Datenschutzrecht mit dem Urheberrecht im Hinblick auf die rechtssystematische Einordnung, die historische Entwicklung und die aktuelle praktische Bedeutung, so zeigt sich, daß es kaum größere Gegensätze zwischen zwei Rechtsmaterien geben könnte. Beim Datenschutzrecht, das dem Besonderen Verwaltungsrecht zuzuordnen ist, handelt es sich um ein erst in der zweiten Hälfte dieses Jahrhunderts entwickeltes Rechtsgebiet.¹ Obwohl das Datenschutzrecht aufgrund seines Regelungsumfanges beinahe jeden Staatsbürger in irgendeiner Form betrifft, fehlt weitgehend ein entsprechendes Rechts- bzw. Unrechtsbewußtsein für Datenschutzangelegenheiten. Vielfach werden die Vorschriften des Datenschutzgesetzes aus Unwissenheit einfach ignoriert.² Ganz anders ist die Situation beim Urheberrecht, welches ein sogenanntes Sonderprivatrecht darstellt. Hier sind Rechtsverletzungen seltener auf mangelndes Unrechtsbewußtsein, sondern eher auf wirtschaftliche Überlegungen zurückzuführen. Das mag vielleicht daran liegen, daß es sich beim Urheberrecht um einen schon lange geregelten Rechtsbereich handelt.³ Wobei zu betonen ist, daß die Gefahr der Entdeckung und Sanktionierung im Falle urheberrechtlicher Verstöße weitaus größer ist als bei Verletzungen der Datenschutzbestimmungen.

Trotz aller Gegensätze in Vergangenheit und Gegenwart lassen sich Datenschutzrecht und Urheberrecht unter dem Oberbegriff des Informationsrechts zusammenfassen. Sämtliche Bereiche dieses Informationsrechts⁴ haben eine vom Wandel der Gesellschaft zu einer Informationsgesellschaft geprägte gemeinsame Zukunft. Ein Kennzeichen der Informationsgesellschaft ist der Umstand, daß deren Mitglieder ihre Informationen nicht mehr nur über die traditionellen Medien, sondern auch mit Hilfe neuer Informations- und Kommunikationstechniken beziehen. Diese neuen "Techniken", unter denen das Internet eine herausragende Stellung einnimmt, schaffen die Möglichkeit, daß sehr viele

¹ Das erste Datenschutzgesetz für Deutschland (BRD) trat 1977 in Kraft, während die erste datenschutzrechtliche Regelung für Österreich 1980 wirksam wurde.

² Vgl. Kapitel 2.3

³ Urheberrechtliche Regelungen gibt es im deutschsprachigen Raum bereits seit dem 13. Jahrhundert ("Bücherfluch" des Sachsenspiegels).

⁴ Darunter lassen sich auch noch zahlreiche andere Rechtsgebiete, wie etwa das Medienrecht, das Telekommunikationsrecht oder das Computerstrafrecht, subsumieren.

Menschen sehr viele Informationen sehr schnell bekommen. Durch diesen Informationszuwachs wird die Gesellschaft insofern grundlegend verändert, als Informationen eine immer größere Bedeutung erhalten (vgl. Lanfermann 1998, 1f.). Ein Mangel an Information birgt die Gefahr einer finanziellen und sozialen Schlechterstellung in sich. Allerdings stehen Informationsbereitstellung und Informationsgewinnung nicht im rechtsfreien Raum. Vor allem die Bestimmungen des Datenschutz- und des Urheberrechts sind stets zu beachten, wobei die aktuellen Gesetzesfassungen dieser zwei Rechtsbereiche sowohl in Österreich als auch in Deutschland der technischen und gesellschaftlichen Entwicklung der letzten Jahre nicht immer Rechnung tragen; in manchen Bereichen besteht ein dringender Anpassungsbedarf.

2 Schutz von Daten

2.1 Rechtslage in Österreich

Der Schutz von Daten wird in Österreich durch das Österreichische Datenschutzgesetz (ÖDSG) aus dem Jahre 1978⁵ geregelt, welches mit 1. Jänner 1980 in Kraft trat. Das ÖDSG wurde unter dem Eindruck der Möglichkeiten der (damals) modernen Datenverarbeitung geschaffen, um die Persönlichkeitssphäre der Bürger zu schützen (vgl. Singer 1991, 1). Es stellt eine umfassende Regelung hinsichtlich der Behandlung personenbezogener Daten dar.

Ein wesentliches Merkmal des ÖDSG ist die Aufteilung dieses Gesetzes in Bestimmungen, die ausschließlich für den öffentlichen Bereich⁶ - also für Datenverarbeitungen durch Bund, Länder, Gemeinden, Kammern, Sozialversicherung, etc. - gelten, in solche, welche ausschließlich für den privaten Bereich - also hauptsächlich für Datenverarbeitungen durch private Unternehmen, Vereine, politische Parteien oder Religionsgemeinschaften - gelten und in solche, die für beide Bereiche gelten. Abbildung 1 zeigt neben der Gliederung des ÖDSG diese Aufteilung. Darüber hinaus werden beinahe sämtliche Rechte der von der jeweiligen Datenverarbeitung Betroffenen auf das Vorliegen einer automationsunterstützten⁷ Datenverarbeitung beschränkt. Lediglich der Anspruch des einzelnen auf Geheimhaltung der ihn betreffenden personenbezogenen Daten, soweit er daran ein schutzwürdiges Interesse hat, erstreckt sich auch auf manuelle Datenverarbeitungen.⁸ Theoretisch bestünde demnach die Möglichkeit, Teile der Datenverarbeitung auf "händischen Betrieb" umzustellen, um den meisten Vorschriften des Datenschutzgesetzes zu entgehen.

⁵ Bundesgesetz vom 18. Oktober 1978 über den Schutz personenbezogener Daten, BGBl. 565/1978

⁶ Vgl. §§ 4 und 5 ÖDSG

⁷ Gem. § 3 Z. 5 ÖDSG liegt eine automationsunterstützte Datenverarbeitung vor, wenn diese maschinell und programmgesteuert erfolgt.

⁸ Vgl. § 1 Abs. 1 ÖDSG

Abschnitt	§§	Inhalt	Gültig für öffentlichen Bereich	Gültig für privaten Bereich
-	1-2	Grundrecht auf Datenschutz	+	+
1	3-5	Allgemeine Bestimmungen	+	+
2	6-16	Öffentlicher Bereich	+	
3	17-31	Privater Bereich		+
4	32-34	Internationaler Datenverkehr	+	+
5	35-47	Datenschutzinstitutionen	+	+
6	48-50	Strafbestimmungen	+	+
7	51-59	Übergangs- und Schlußbestimmungen	+	+

Abb. 1: Gliederung des ÖDSG (vgl. Dohr 1996, 84)

Das Grundrecht auf Datenschutz normiert neben dem eben erwähnten Anspruch auf Geheimhaltung die Rechte auf Auskunft, Richtigstellung und Löschung.⁹ Im ersten Abschnitt des ÖDSG werden u.a. die vom Gesetz erfaßten Akteure definiert: der Betroffene (einer Datenverarbeitung) ist demnach derjenige, dessen Daten verwendet werden, wogegen der Auftraggeber derjenige ist, der die Daten des Betroffenen entweder selbst oder unter Heranziehung eines sogenannten Dienstleisters automationsunterstützt verarbeitet.¹⁰ Den Kern des ÖDSG bilden der 2. und 3. Abschnitt, wo, jeweils für den öffentlichen und den privaten Bereich getrennt, einerseits die Rechte der Betroffenen auf Auskunft, Richtigstellung und Löschung näher ausgestaltet sind und andererseits die jeweiligen Voraussetzungen für die Zulässigkeit einer bestimmten Datenverarbeitung einschließlich Registrierungspflichten festgelegt sind. Im 4. Abschnitt des Gesetzes wird geregelt, unter welchen Voraussetzungen eine Datenübermittlung ins Ausland genehmigungspflichtig ist. Eine solche Genehmigung erfolgt durch die Datenschutzkommission, die gemeinsam mit dem Datenschutzrat und dem Datenverarbeitungsregister den Inhalt des 5. Abschnittes bildet. Die im 6. Abschnitt genannten Justiz- und Verwaltungsstrafbestimmungen haben nur geringe praktische Bedeutung.

Eine der zentralen Bestimmungen des ÖDSG ist sicher das Recht auf Auskunft¹¹, also das Recht des von der Datenverarbeitung Betroffenen auf Bekanntgabe der zu seiner Person bei der auskunftspflichtigen Stelle gespeicherten Daten. Im einzelnen hat der Betroffene Anspruch auf Auskunft darüber, wer

⁹ Während der Geheimhaltungsanspruch in § 1 Abs. 1 ÖDSG abschließend geregelt ist, werden die Rechte auf Auskunft, Richtigstellung und Löschung im § 1 Abs. 3, 4 und 5 ÖDSG lediglich in Grundzügen angeführt. Die nähere Ausgestaltung dieser Rechte erfolgt erst im 2. (§§ 11 und 12 ÖDSG) bzw. 3. Abschnitt (§§ 25-27 ÖDSG) des ÖDSG.

¹⁰ Vgl. § 3 Z. 2-4 ÖDSG

¹¹ Geregelt in den §§ 11 bzw. 25 ÖDSG

die Daten ermittelt und verarbeitet hat, woher sie stammen, welchen Inhalt sie haben und für welchen Zweck sie verwendet werden. Im öffentlichen Bereich ist noch zusätzlich die Rechtsgrundlage für die Datenverarbeitung bekanntzugeben, im privaten Bereich gegebenenfalls Name und Anschrift eines Dienstleisters. Voraussetzung für eine Auskunftserteilung, die vom Auftraggeber binnen vier Wochen und grundsätzlich kostenlos zu erfolgen hat, ist eine entsprechende Anfrage eines Betroffenen. Im Falle einer Auskunftsverweigerung oder mangelhaften Auskunftserteilung steht dem Berechtigten nach dem ÖDSG bezüglich des öffentlichen Bereiches eine Beschwerde an die Datenschutzkommission offen, für Angelegenheiten des privaten Bereiches bleibt lediglich der risikante Gerichtsweg.

2.2 Rechtsvergleich zwischen Österreich und Deutschland

Das aktuelle¹² Bundesdeutsche Datenschutzgesetz (BDSG) trat mit 1. Juni 1991 in Kraft.¹³ Das BDSG bezieht sich, wie das ÖDSG, nur auf personenbezogene Daten und soll ebenfalls nicht die Daten als solche schützen, sondern die Persönlichkeit desjenigen, auf den sich die Daten beziehen (vgl. Däubler 1996, 21). Abbildung 2 gibt einen Überblick über die Gliederung des BDSG, aus der sich - als weitere Gemeinsamkeit mit dem ÖDSG - ebenfalls eine Einschränkung der Gültigkeit mancher Bestimmungen auf den öffentlichen bzw. privaten¹⁴ Bereich ergibt. Hinsichtlich Aufbau und Inhalt gibt es zwar zahlreiche weitere Parallelen zwischen ÖDSG und BDSG¹⁵, aber doch auch einige gravierende Unterschiede.

<u>Abschnitt</u>	<u>§§</u>	<u>Inhalt</u>	<u>Gültig für öffentl. Bereich</u>	<u>Gültig für privaten Bereich</u>
1	1-11	Allgemeine Bestimmungen	+	+
2	12-26	Datenverarbeitung der öffentlichen Stellen	+	
3	27-38	Datenverarbeitung nicht-öffentlicher Stellen		+
4	39-42	Sondervorschriften	+	+
5	43-44	Schlußvorschriften	+	+

Abb. 2: Gliederung des BDSG (vgl. Herbek/Schillinger 1996, 119)

¹² Davor gab es aber bereits eine bundesweite datenschutzrechtliche Regelung, das BDSG vom 27. Jänner 1977 (BGBl. I, 201).

¹³ BGBl. I, 2954 vom 20.12.1990

¹⁴ Die Bezeichnungen "privat" im ÖDSG und "nicht-öffentlich" im BDSG werden deklungs-gleich verwendet.

¹⁵ Auf diese soll aber im Rahmen dieses Beitrages nicht näher eingegangen werden.

Die im ÖDSG in eigenen Abschnitten geregelten Bereiche des internationalen Datenverkehrs und der Datenschutzinstitutionen sind im BDSG in den 2. bzw. 3. Abschnitt integriert. Dafür enthält das BDSG einen eigenen Abschnitt mit Sonderbestimmungen für bestimmte Bereiche, wie etwa Medien oder wissenschaftliche Forschung. Die Strafbestimmungen sind im BDSG in den Schlußvorschriften normiert.

Ein wesentlicher Unterschied zum ÖDSG liegt in der Einbeziehung manuell erstellter Datensammlungen ins BDSG, wobei allerdings Akten¹⁶ grundsätzlich nicht erfaßt werden. Weiters normiert das BDSG mit den Rechten auf Sperrung und Information, die dort neben den Rechten auf Auskunft, Richtigstellung und Löschung angeführt sind, weitreichendere Befugnisse der Betroffenen. Eine Sperrung von Daten ist u.a. dann vorgesehen¹⁷, wenn deren Richtigkeit vom Betroffenen bestritten wird und sich weder die Richtigkeit noch die Unrichtigkeit feststellen läßt. Aufgrund des nur für den nicht-öffentlichen Bereich eingeführten Informationsrechts¹⁸ ist der Betroffene vom Auftraggeber¹⁹ von der erstmaligen Speicherung personenbezogener Daten und von der Art der gespeicherten Daten in Kenntnis zu setzen. Durch diese Basisinformation soll der Betroffene in die Lage versetzt werden, sich bei Interesse mit Hilfe seines Auskunftsrechtes genauer zu informieren (vgl. Däubler 1996, 243).

Wenig Gemeinsamkeiten existieren zwischen ÖDSG und BDSG auch bezüglich der eingerichteten Kontrollinstanzen für den Datenschutz. Während es sich beim Bundesbeauftragten für den Datenschutz nach BDSG noch um eine der Datenschutzkommission nach ÖDSG ähnliche Kontrollinstitution für den öffentlichen Bereich handelt, sind im BDSG mit den betrieblichen Datenschutzbeauftragten und mit der Aufsichtsbehörde auch für den privaten Bereich außergerichtliche Kontrollinstanzen gesetzlich verankert. Es handelt sich hierbei um ein zweistufiges Kontrollsystem, bei dem der betriebliche Datenschutzbeauftragte die interne Instanz bildet, dessen Tätigkeit durch die Aktivitäten der Aufsichtsbehörde als externer Instanz ergänzt werden. Sämtliche nicht-öffentlichen Stellen, die personenbezogene Daten verarbeiten und mindestens zwanzig²⁰ Mitarbeiter beschäftigen, haben einen betrieblichen Datenschutzbeauftragten zu bestellen.²¹ Somit sind Betroffene, die sich in ihren Datenschutzrechten als verletzt erachten, aufgrund der Bestimmungen des BDSG zur Durchsetzung ihrer Befugnisse weder im öffentlichen noch im privaten Bereich auf den Rechtsweg angewiesen.

¹⁶ Vgl. § 3 Abs. 2 und 3 BDSG

¹⁷ Gem. § 20 Abs. 4 BDSG (öffentlicher Bereich) bzw. § 35 Abs. 4 BDSG (nicht-öffentlicher Bereich)

¹⁸ Gem. § 33 BDSG

¹⁹ Dafür wird im BDSG der Begriff "Speichernde Stelle" verwendet.

²⁰ Im Falle einer automationsunterstützten Datenverarbeitung tritt diese Verpflichtung bereits ab fünf Beschäftigten ein.

²¹ Gem. § 36 Abs. 1 BDSG

2.3 Probleme und Anpassungsbedarf

Sowohl das österreichische als auch das deutsche Datenschutzrecht weisen aus heutiger Sicht eine Reihe von Normierungsdefiziten auf, sodaß für manches aktuelle datenschutzrechtliche Problem keine gesetzliche Regelung existiert. Ursachen für diese Defizite sind u.a. der Wandel des Gefährdungspotentials der Datenverarbeitung und die Veränderung des gebotenen Schutzzweckes datenschutzrechtlicher Bestimmungen. Während bei der Konzeption der betrachteten Datenschutzgesetze noch von einer Dominanz der Großcomputer und einer zukünftigen Konzentration beinahe sämtlicher Daten in einigen wenigen nationalen Datenbanken ausgegangen wurde, gibt es mittlerweile Mikrocomputer mit einer wesentlich höheren Speicherkapazität als jener der Großcomputer der 70-er Jahre, deren Vernetzung zu immer größeren Informationspools führt. Eine staatliche bzw. multinationale Aufsicht über all diese Daten(verarbeitungen) ist praktisch nicht mehr realisierbar. Zudem setzte sich in den 80-er Jahren die Erkenntnis durch, daß der Schutzzweck des Datenschutzes nicht der umfassende aktive staatliche Schutz des einzelnen vor Mißbrauch der ihn betreffenden Daten sein müsse, sondern eher die Garantie einer weitgehenden informationellen Selbstbestimmung. Nicht der Staat, sondern der einzelne ist über die Verarbeitung personenbezogener Daten zu informieren; dieser kann dann selbst bestimmen, ob er die jeweilige Verarbeitung akzeptiert oder unterbinden möchte. Erst im Falle einer Mißachtung dieser Rechte des einzelnen soll der Staat unterstützend zur Seite stehen (vgl. Mayer-Schönberger 1998, 245f.). Zusätzlich tritt noch das Problem der Internationalität moderner Informations- und Kommunikationsnetze auf. Der nationale Datenschutz ist prinzipiell auf das Territorium eines Staates beschränkt, wogegen es für die Daten auf diesen Netzen keinerlei Grenzen gibt. Somit besteht die Gefahr der Umgehung des Datenschutzes durch Transfer der Daten zum Zwecke der Weiterverarbeitung in Länder mit keinen²² oder schwächeren Datenschutznormen (vgl. Mayer-Schönberger 1997, 165).

Den Anforderungen an ein - zumindest europaweit - einheitliches Datenschutzrecht unter Berücksichtigung der Bedürfnisse der globalen Informationsgesellschaft soll vor allem durch die Umsetzung der EU-Datenschutzrichtlinie aus dem Jahre 1995²³ entsprochen werden. Diese Richtlinie soll zu einer Harmonisierung der Datenschutzbestimmungen innerhalb der EU auf hohem Niveau und zu einer Vereinheitlichung des Umganges mit Datenübermittlungen ins EU-Ausland führen. Die folgenden Ausführungen dienen der Darstellung der wichtigsten Inhalte der Richtlinie, die bis zum 24. Oktober 1998 in sämtlichen EU-Staaten in nationales Recht umzusetzen sind.²⁴

²² So existiert etwa in den USA, einem der wesentlichen Akteure in den globalen Informationsnetzen, bislang kein umfassendes Datenschutzrecht.

²³ Richtlinie 95/46/EG des Europäischen Parlaments und des Rates vom 24. Oktober 1995 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr

²⁴ Dieser Termin wird allerdings voraussichtlich von einigen EU-Staaten, darunter auch Österreich und Deutschland, nicht eingehalten werden.

Im Gegensatz zum ÖDSG beschränkt sich die Richtlinie auf den Schutz natürlicher Personen.²⁵ Dafür gilt die Richtlinie sowohl für die ganz oder teilweise automatisierte Verarbeitung personenbezogener Daten als auch für die nicht automatisierte Verarbeitung personenbezogener Daten. Weiters wird die in den meisten nationalen Datenschutzgesetzen vorhandene Dreiteilung in Bestimmungen für den öffentlichen Bereich, Bestimmungen für den privaten Bereich und Bestimmungen für beide Bereiche aufgegeben, indem alle Bestimmungen der Richtlinie für beide Bereiche gelten. Vom Geltungsbereich ausgenommen sind lediglich einige spezielle Formen der Datenverarbeitung, wie etwa jene der Geheimdienste oder jene im ausschließlich familiären Bereich. Eine - zumindest für Österreich - gänzliche Neuerung stellt die Normierung einer Informationspflicht des Auftraggebers gegenüber der betroffenen Person bei der Erhebung personenbezogener Daten dar. Diese Informationspflicht stellt einen Eckpfeiler des Rechts auf informationelle Selbstbestimmung der Betroffenen dar. Die Information über Verarbeitung personenbezogener Daten wird somit von einer Holschuld des Betroffenen im Falle eines reinen Auskunftsrechtes zu einer Bringschuld des Auftraggebers. Einschränkungen der Informationspflicht sind nur in wenigen Fällen vorgesehen, wie etwa bei möglichen Gefährdungen der öffentlichen Sicherheit. Im Zusammenhang mit der Kontrolle der nationalen Datenschutzvorschriften sieht die Richtlinie die Einrichtung von unabhängigen Kontrollstellen vor, wie sie in Deutschland mit dem Bundesbeauftragten für den Datenschutz im öffentlichen Bereich und der Aufsichtsbehörde im privaten Bereich bereits existieren. Zudem muß jede Person bei Verletzung der in der Richtlinie garantierten Rechte einen Rechtsbehelf bei Gericht einlegen können. Für den internationalen Datenverkehr legt die Richtlinie fest, daß ein Datentransfer in Drittländer²⁶ nur stattfinden darf, wenn das Drittland ein angemessenes Schutzniveau²⁷ gewährleistet.

Aufgrund dieser Richtlinie wären im ÖDSG bzw. BDSG unter anderem die in Abbildung 3 enthaltenen Anpassungen notwendig. Derzeit liegen zwar schon umfassende Entwürfe für neue Datenschutzgesetze in Österreich²⁸ und Deutschland²⁹ vor, doch dürfte es sich dabei noch keineswegs um die endgültigen Versionen handeln.³⁰ Aus diesem Grund soll auch eine nähere Betrachtung dieser Entwürfe unterbleiben. Erst anhand der im jeweiligen Bundesgesetzblatt verlautbarten Version des neuen Datenschutzgesetzes wird man die Übereinstimmung dieses Gesetzes mit den Vorgaben der EU-Richtlinie überprüfen können.

²⁵ Einen Schutz von Daten juristischer Personen gibt es innerhalb der EU außer in Österreich nur in Dänemark und Luxemburg.

²⁶ Das sind alle Nicht-EU-Mitgliedsstaaten.

²⁷ Wobei die exakte Definition der Angemessenheit ein eigenes Problem darstellt.

²⁸ Der offizielle Entwurf eines Bundesgesetzes über den Schutz personenbezogener Daten (DSG 1998) wurde im März 1998 vorgestellt und mittlerweile bereits mehrfach überarbeitet.

²⁹ Der Referentenentwurf der Bundesregierung zur Novellierung des Bundesdatenschutzgesetzes stammt bereits aus dem Dezember 1997.

³⁰ Laut Auskunft durch das Österreichische Bundeskanzleramt, das in Österreich für die Datenschutzgesetzgebung zuständig ist.

<u>Bereich</u>	<u>Anpassungsbedarf im ÖDSG</u>	<u>Anpassungsbedarf im BDSG</u>
Einbeziehung der manuellen Daten(verarbeitung)	ja ³¹	nein
Keine Trennung zwischen öffentlichem und privatem Bereich	ja	ja
Informationspflicht	ja	ja ³²
Einrichtung unabhängiger Kontrollstellen	ja ³³	nein
Ordentlicher Rechtsweg	ja ³⁴	ja
Datentransfer in Drittländer	ja ³⁵	ja

*Abb. 3: Anpassungsbedarf aufgrund der EU-Richtlinie
(vgl. Mayer-Schönberger 1998, 247ff. ; Runge 1998, 46ff.)*

Angesichts der zahlreichen erforderlichen Modifikationen im Datenschutzbereich sollten neben der gesetzestechnischen Ausgestaltung dieser Änderungen auch Maßnahmen zur Hebung des Bewußtseins für Datenschutzangelegenheiten sowohl auf seiten der Auftraggeber als auch auf seiten der Betroffenen überlegt werden, um ein Dasein des Datenschutzrechts als praktisch totes Recht zu verhindern. Eine derartige Gefahr bestand und besteht nicht nur für datenschutzrechtliche Spezialbestimmungen, wie etwa die im ÖDSG vorgeschriebenen Datensicherheitsmaßnahmen, sondern auch für Kernbereiche des Datenschutzrechts, zu denen sicherlich das Auskunftsrecht zu zählen ist. Gespräche des Autors mit Datenschutzexperten führten zur Vermutung, daß dieses Recht von den Auskunftspflichtigen in der Praxis oftmals einfach ignoriert wird. Am Institut für Informationswissenschaft in Graz wurden deshalb konkrete

³¹ Im Falle dieser Einbeziehung wird bei jenen Auftraggebern ein hoher Organisationsbedarf entstehen, die bisher auf manuelle Datenverarbeitungen zurückgegriffen haben, um nicht den Bestimmungen des ÖDSG unterworfen zu sein (vgl. Dohr/Schillinger 1998, 156).

³² Die Informationspflicht des Auftraggebers ist auf den öffentlichen Bereich zu erweitern.

³³ Anpassungsbedarf besteht nur für den privaten Bereich. Für den öffentlichen Bereich existiert bereits die Datenschutzkommission.

³⁴ Auch für den öffentlichen Bereich ist die Möglichkeit vorzusehen, bei Verletzung der Datenschutzrechte einen Rechtsbehelf bei Gericht einlegen zu können.

³⁵ Hier sind nur geringe Anpassungen notwendig.

Datenschutzanfragen³⁶ durchgeführt, um diese Vermutung entweder zu bestätigen oder zu widerlegen. Das Ergebnis dieses Experiments war eindeutig negativ. Von 19 Datenschutzanfragen wurden 10 überhaupt nicht bzw. absolut unbrauchbar beantwortet, 4 Auftraggeber gaben eine etwas mangelhafte Auskunft, und nur 5 Auftraggeber antworteten im Sinne des ÖDSG. Besonders bedenklich erscheint der Umstand, daß keiner von jenen 6 Auftraggebern, die dem öffentlichen Bereich zuzuordnen waren, die Datenschutzanfrage ordnungsgemäß beantwortete. Wobei sicherlich davon auszugehen ist, daß mancher Fehler auf mangelnde Erfahrung der Auftraggeber mit Datenschutzanfragen zurückzuführen war. Diese mangelnde Erfahrung ist wiederum ein Resultat der kaum in Anspruch genommenen Datenschutzrechte der Betroffenen. Es fehlt hier meist das Wissen um diese Rechte. Dabei gäbe es viele Situationen, in denen eine gesetzeskonform gestellte und beantwortete Datenschutzanfrage für den Betroffenen äußerst informativ und nützlich wäre.

3 Schutz von Wissen

3.1 Rechtslage in Österreich

Das Österreichische Urheberrechtsgesetz (öUrhG) stammt aus dem Jahre 1936³⁷, wobei die ursprüngliche Kodifizierung immer wieder aktuellen Entwicklungen angepaßt wurde. Es schützt in erster Linie die Urheber von Werken in persönlichkeitsrechtlicher und wirtschaftlicher Hinsicht. Abbildung 4 zeigt die Systematik des öUrhG. Kernstück dieses Gesetzes ist das 1. Hauptstück, das auch als Urheberrecht im engeren Sinn bezeichnet wird (vgl. Kucsko 1996, 19). Schutzobjekt sind hierbei Werke. Die sogenannten verwandten Schutzrechte³⁸ schützen dagegen u.a. Vorträge, Aufführungen, Lichtbilder und Briefe. Hinsichtlich der Rechtsdurchsetzung ist zwischen zivilrechtlichen Ansprüchen des Urhebers, wie etwa dem Unterlassungsanspruch, und strafrechtlichen Sanktionen für den Verletzer von Urheberrechten zu unterscheiden. Für den Anwendungsbereich des öUrhG gelten das Prinzip der Anknüpfung an den Erscheinungsort und das Staatsbürgerschaftsprinzip (vgl. Ciresa 1997, 41f.). Demnach genießen alle in Österreich erschienenen³⁹ Werke - unabhängig von der Staatsangehörigkeit des Urhebers - urheberrechtlichen Schutz im Inland. Zusätzlich sind hier auch sämtliche im Ausland erschienenen Werke geschützt, sofern der Urheber österreichischer Staatsbürger ist.

³⁶ Im Rahmen eines Proseminars am Institut wurden Studenten im Sommersemester 1998 beauftragt, Datenschutzanfragen an jene öffentlichen Stellen bzw. privaten Unternehmen zu richten, wo eventuell Daten über sie gespeichert sein könnten.

³⁷ BGBl. 111/1936

³⁸ Die verwandten Schutzrechte und das Urheberrecht im engeren Sinn bilden zusammen das Urheberrecht im weiteren Sinn.

³⁹ Ein Werk gilt gem. § 9 Abs. 1 öUrhG dann als erschienen, sobald es mit Einwilligung des Urhebers der Öffentlichkeit dadurch zugänglich gemacht worden ist, daß Werkstücke in genügender Anzahl feilgehalten oder in Verkehr gebracht worden sind. Somit ist ein Werk etwa dann erschienen, wenn es als Buch auf den Markt kommt.

<u>Hauptstück</u>	<u>§§</u>	<u>Inhalt</u>
1	1-65	Urheberrecht an Werken der Literatur und der Kunst
2	66-80	Verwandte Schutzrechte
3	81-93	Rechtsdurchsetzung
4	94-100	Anwendungsbereich des Gesetzes
5	101-114	Übergangs- und Schlußbestimmungen

Abb. 4: Gliederung des öUrHG

Zur Erlangung des urheberrechtlichen Schutzes bedarf es keines Formalaktes; es genügt bereits die Schaffung eines Werkes, also einer eigentümlichen geistigen Schöpfung⁴⁰, durch einen Urheber⁴¹. Während sich der Urheberbegriff als weitgehend zeitlos erwiesen hat, mußte der Werkbegriff immer wieder angepaßt werden, wobei seit einer Anpassung im Jahre 1993⁴² Computerprogramme per Gesetz zu den Werken der Literatur zählen.⁴³ Inhaltlich gewährt das Urheberrecht dem Inhaber das ausschließliche Recht, sein Werk wirtschaftlich zu nutzen⁴⁴ sowie das Recht auf Schutz seiner geistigen Interessen am Werk⁴⁵. Innerhalb der Verwertungsrechte ist neben dem Verbreitungsrecht sicherlich das Vervielfältigungsrecht von besonderer Bedeutung, von dem allerdings die Vervielfältigung von Werken zum eigenen Gebrauch bzw. zum Hochschulgebrauch nicht erfaßt wird, da es sich dabei um freie Werknutzungen handelt. Die freie Werknutzung für Unterrichtszwecke an (Hoch)schulen wurde im Rahmen der Urheberrechtsgesetznovelle 1996⁴⁶ dahingehend erweitert, daß die für den Unterrichtszweck erforderlichen Vervielfältigungsstücke für eine ganze Schulklasse bzw. Lehrveranstaltung in der notwendigen Zahl hergestellt werden dürfen (vgl. Gamerith 1997, 103). Das Urheberrecht ist vererblich und endet erst 70 Jahre nach dem Tod des Urhebers. Im Gegensatz dazu ist es unübertragbar; es können lediglich sogenannte Werknutzungsbewilligungen in Form von Urheberrechtsverträgen eingeräumt werden, deren häufigste Erscheinungsformen Verlagsverträge zwischen Urhebern und Verlegern und Wahrnehmungsverträge zwischen Urhebern und Verwertungsgesellschaften sind.

⁴⁰ Vgl. § 1 öUrHG

⁴¹ Gem. § 10 Abs. 1 öUrHG ist derjenige Urheber eines Werkes, der dieses geschaffen hat.

⁴² BGBl. 93/1993

⁴³ Vgl. § 2 Z. 1 öUrHG

⁴⁴ Man spricht hier von den sogenannten Verwertungsrechten.

⁴⁵ Hier wird auch der Begriff "Urheberpersönlichkeitsrecht" verwendet. Zu diesem Recht zählen das Veröffentlichungsrecht, der Schutz der Urheberschaft, das Recht auf Wahl der Urheberbezeichnung und der Werkschutz.

⁴⁶ BGBl. 151/1996

3.2 Rechtsvergleich zwischen Österreich und Deutschland

Ein Vergleich des Deutschen Urheberrechtsgesetzes (dUrHG) aus dem Jahre 1965⁴⁷ mit dem öUrHG zeigt die weitgehende Übereinstimmung dieser beiden Gesetze. Unterschiede finden sich lediglich hinsichtlich der Gesetzesgliederung, wie aus Abbildung 5 ersichtlich wird, und hinsichtlich der Lösung von Detailproblemen. Als Beispiel für einen inhaltlichen Unterschied kann etwa die nach dem dUrHG gebotene Beschränkung der Vervielfältigung von Werken zum Schulgebrauch auf Schulen exklusive Hochschulen angeführt werden, wobei auch Schulen nicht zur Herstellung von Vervielfältigungsstücken eines gesamten Werkes befugt sind, sondern lediglich zur Erzeugung von Vervielfältigungsstücken von kleinen Teilen eines Werkes.⁴⁸ Hochschulen dürfen derartige Vervielfältigungsstücke von Werkteilen ausschließlich für Prüfungszwecke herstellen (vgl. Hillig 1995, XXIV).

Teil	§§	Inhalt
1	1-69g	Urheberrecht ⁴⁹
2	70-87	Verwandte Schutzrechte ⁵⁰
3	88-95	Besondere Bestimmungen für Filme ⁵¹
4	96-119	Gemeinsame Bestimmungen für 1. und 2. Teil ⁵²
5	120-143	Anwendungsbereich, Übergangs- und Schlußbestimmungen ⁵³

Abb. 5: Gliederung des dUrHG

3.3 Probleme und Anpassungsbedarf

In der postindustriellen Informationsgesellschaft nehmen geistige Leistungen eine zentrale Stellung ein. Das traditionelle Urheberrecht, das natürlich lange vor der Entstehung und Verbreitung der "neuen" Medien geschaffen wurde, ist diesen neuen Herausforderungen anzupassen, um den erwünschten Schutz zu bieten. Hinsichtlich der Einordnung und des Schutzes von Computerprogrammen ist zu betonen, daß diese als neue Werkart sowohl ins dUrHG⁵⁴ als auch

⁴⁷ BGBl. I S. 1273/1965

⁴⁸ Vgl. § 53 Abs. 2 und 3 dUrHG

⁴⁹ Dieser Teil ist weitgehend identisch mit dem 1. Hauptstück des öUrHG.

⁵⁰ Entspricht größtenteils dem 2. Hauptstück des öUrHG.

⁵¹ Im öUrHG sind ähnliche Bestimmungen im 1. Hauptstück enthalten.

⁵² Hier ist in erster Linie die Rechtsdurchsetzung geregelt, die sich im öUrHG im 3. Hauptstück findet.

⁵³ In diesem Teil sind jene Bestimmungen zusammengefaßt, die im öUrHG im 4. und 5. Hauptstück geregelt sind.

⁵⁴ Gem. § 2 Abs. 1 Z. 1 dUrHG

ins öUrhG⁵⁵ Eingang fanden. In Österreich erfolgte diese Integration allerdings erst im Jahre 1993 im Zuge der verpflichtenden Umsetzung der EU-Richtlinie über den Rechtsschutz von Computerprogrammen⁵⁶. Im dUrhG kam es aufgrund dieser Richtlinie hingegen zu einer Herabsetzung⁵⁷ der Voraussetzungen für die Schutzfähigkeit von Computerprogrammen (vgl. Hertin 1996, 102).

Die urheberrechtlichen Probleme, die sich aus der exponentiell wachsenden Nutzung des Internets als Informations- und Kommunikationsmedium ergeben, sind noch keinesfalls umfassend gelöst. Ausgangspunkt dieser Problematik ist die Digitalisierung von urheberrechtlich geschützten Werken. Durch eine solche Digitalisierung wird die Verletzlichkeit von geistigem Eigentum wesentlich erhöht, denn das Kopieren eines digitalisierten Werkes auf die eigene Festplatte bzw. Diskette ist noch wesentlich einfacher und unauffälliger als eine Vervielfältigung eines physischen Werkstückes mittels reprographischer Verfahren. Abhilfe könnte hierbei die Einführung eines digitalen Wasserzeichens, verbunden mit entsprechenden Vergütungsansprüchen des Urhebers, schaffen. Eine realistische Gesamtlösung wäre die Kombination des digitalen Wasserzeichens zur Wahrung der Urheberpersönlichkeitsrechte mit einer Gerätevergütung⁵⁸ zur Wahrung der wirtschaftlichen Verwertungsrechte des Urhebers, wobei sich im Falle einer solchen pauschalierten Vergütung die Frage nach einer gerechten Lösung für Vervielfältigungen zum eigenen Gebrauch stellen würde (vgl. Lehmann 1997, 28ff.; Herrigel 1998).

Ein weiteres urheberrechtliches Problemfeld ergibt sich aus der Internationalität der Informationsflüsse im Internet. Die nationalen Urheberrechtsgesetze sind in der Regel durch das Territorialitätsprinzip geprägt, was zur Folge hat, daß sich die urheberrechtlichen Ansprüche nach dem Recht des Staates richten, in dem eine Benützungs- oder Verletzungshandlung gesetzt wird. Da die nationalen Schutzniveaus stark differieren, gewinnt das Instrumentarium mehrseitiger Staatsverträge zunehmend an Bedeutung. Besonders erwähnenswert ist in diesem Zusammenhang die Revidierte Berner Übereinkunft zum Schutz von Werken der Literatur und Kunst, der seit 1989 mit dem Beitritt der USA praktisch alle westlichen "Industrienationen" angehören⁵⁹. Die Berner Übereinkunft normiert einerseits den Grundsatz der Inländerbehandlung⁶⁰ und legt andererseits für alle Verbandsstaaten einen urheberrechtlichen Mindestschutz⁶¹ fest. Für den

⁵⁵ Vgl. 3.1.

⁵⁶ Richtlinie 91/250/EWG des Rates vom 14.5.1991 über den Rechtsschutz von Computer-Programmen

⁵⁷ Anstelle einer gewissen Schöpfungshöhe ist jetzt nur noch die Individualität der Software maßgebend.

⁵⁸ Ähnlich der Reprographievergütung bei Kopiergeräten

⁵⁹ Ende 1995 gehörten der Berner Übereinkunft bereits 117 Staaten an.

⁶⁰ Demnach genießen Urheber, die Staatsbürger eines Mitgliedsstaates der Berner Übereinkunft sind, in jedem Mitgliedsstaat den urheberrechtlichen Schutz inländischer Urheber.

⁶¹ So wird dort z.B. eine Mindestschutzfrist von 50 Jahren (ab dem Tod des Urhebers) festgelegt.

Bereich der EU existiert darüber hinaus ein Richtlinienentwurf⁶² zur Harmonisierung des Urheberrechts und der verwandten Schutzrechte in der Informationsgesellschaft. Der Entwurf ist u.a. den Themen "Harmonisierung des Verbreitungsrechts" und "Harmonisierung des Vervielfältigungsrechts" gewidmet (vgl. Haller 1998).

4 Resümee

Insgesamt läßt sich festhalten, daß weder Datenschutz- noch Urheberrecht mit den aktuellen technischen und gesellschaftlichen Entwicklungen Schritt halten können. Trotz zahlreicher Novellierungen der entsprechenden nationalen Vorschriften in den vergangenen zehn Jahren hat sich die Kluft zwischen Rechts Theorie und Verhaltenspraxis eher vergrößert. Ein effektiver Weg zur Lösung dieses Problems, der mit der zunehmenden Vorlage von EU-Richtlinien und dem wachsenden Abschluß internationaler Übereinkommen bereits eingeschlagen wurde, dürfte in einer zumindest teilweisen Europäisierung bzw. Globalisierung des Rechts liegen.

Literatur

(Ciresa 1997)

Ciresa, M.: *Urheberrecht aktuell*, Wien : Wirtschaftsverlag Carl Ueberreuter 1997

(Däubler 1996)

Däubler, W. / Klebe, T. / Wedde, P.: *Bundesdatenschutzgesetz: Basiskommentar mit der neuen EG-Datenschutzrichtlinie*, Köln : Bund-Verlag 1996

(Dohr 1996)

Dohr, W.: *Das Österreichische Datenschutzgesetz*, in: Fleissner, P. / Choc, M. (Hrsg.): *Datensicherheit und Datenschutz*, Innsbruck / Wien : Studien-Verlag 1996, 83 - 115

(Dohr/Schillinger 1996)

Dohr, W. / Schillinger, B.: *Datenschutz in internationalen Organisationen*, in: Fleissner, P. / Choc, M. (Hrsg.): *Datensicherheit und Datenschutz*, Innsbruck / Wien : Studien-Verlag 1996, 139 - 160

(Gamerith 1997)

Gamerith, H.: *Die wichtigsten Änderungen der Urheberrechtsgesetznovelle 1996*, in ÖBl (Österreichische Blätter für gewerblichen Rechtsschutz und Urheberrecht) 1997, 99 - 106

⁶² Die EU-Kommission hat diesen Vorschlag im Dezember 1997 vorgelegt; er wäre bei Akzeptanz bis zum 30.6.2000 in nationales Recht umzusetzen.

(Haller 1998)

Haller, A.: *Zum EG-Richtlinienvorschlag betreffend Urheberrecht in der Informationsgesellschaft*, in: *medien und recht* 2/1998, 61 - 66

(Herbek/Schillinger 1996)

Herbek, M. / Schillinger, B.: *Das Bundesdeutsche Datenschutzgesetz*, in: Fleissner, P. / Choc, M. (Hrsg.): *Datensicherheit und Datenschutz*, Innsbruck / Wien : Studien-Verlag 1996, 117 - 138

(Herrigel 1998)

Herrigel, A.: *Digitale Wasserzeichen als Urheberschutz*, in: *DuD (Datenschutz und Datensicherheit)* 5/1998, 254 - 257

(Hertin 1996)

Hertin, P.: *Urheberrecht*, Freiburg/Berlin : Haufe 1996

(Hillig 1995)

Hillig, H.-P. (Hrsg.): *Urheber- und Verlagsrecht*, München : C.H. Beck 1995

(Kucsko 1996)

Kucsko, G.: *Österreichisches und europäisches Urheberrecht*, 4. Auflage, Wien : Manz 1996

(Lanfermann 1998)

Lanfermann, H.: *Datenschutzgesetzgebung - gesetzliche Rahmenbedingungen einer liberalen Informationsgesellschaft*, in: *RDV (Recht der Datenverarbeitung)* 1/1998, 1 - 5

(Lehmann 1997)

Lehmann, M. (Hrsg.): *Internet- und Multimediarecht*, Stuttgart : Schäffer-Poeschel 1997

(Mayer-Schönberger 1997)

Mayer-Schönberger, V.: *Das Recht am Info-Highway*, Wien : Orac 1997

(Mayer-Schönberger 1998)

Mayer-Schönberger, V. / Zeger, H. / Kronegger, D.: *Auf dem Weg nach Europa: Zur Novellierung des Datenschutzgesetzes*, in: *ÖJZ (Österreichische Juristen Zeitung)* 1998, 244 - 251

(Runge 1998)

Runge, G.: *Bundesdatenschutzgesetz - Was sich ändert und was bleibt*, in: *Office Management* 3/1998, 46 - 48

(Singer 1991)

Singer, Ch.: *Einige Grundbegriffe*, in: Wittmann, H. et al.: *Datenschutzrecht im Unternehmen*, Wien : Verlag Medien und Recht 1991, 1 - 16

Telekommunikation und Partizipation

Hans W. Giessen

Informationswissenschaft,
Universität des Saarlandes,
D-66041 Saarbrücken
gi@cops.uni-sb.de

Zusammenfassung

Mit der Popularisierung neuer Medien wurde regelmäßig die Hoffnung verbunden, daß sie zu einer aufgeklärteren und demokratischeren Gesellschaft beitragen würden - insbesondere dadurch, daß sie erstmals auch in einer Massengesellschaft Partizipation ermöglichten. In Deutschland untersuchten Vertreter der ‚Kritischen Theorie‘ bereits in den zwanziger Jahren die Beziehungen zwischen den (damals wichtigen) Medien und Formen direkter Demokratie (besonders wichtig: Walter Benjamin). Obwohl die damals formulierten Einsichten in den siebziger Jahren noch einmal weit rezipiert wurden (z.B. Hans Magnus Enzensberger), setzte sich inzwischen offenbar die Ansicht durch, daß sie keine befriedigende Erklärung geben können - zumindest nicht hinsichtlich der Medien, auf die sich die entsprechenden Aussagen bezogen hatten; in der Folge verloren sie dementsprechend an Bedeutung. Nun ist bemerkenswert, daß in der Gegenwart und bezogen auf die heute aktuellen beziehungsweise sich neu entwickelnden Formen der Telekommunikation eine Diskussionen geführt wird, die deutliche Ähnlichkeiten zu den Gedanken der ‚Kritischen Theorie‘ aufweist. Aus diesem Grund soll versucht werden, die theoretischen Aussagen der Kritischen Theorie auf die neuen Formen der Telekommunikation anzuwenden; zudem soll untersucht werden, welche Probleme und Argumente der Gültigkeit der Kritischen Theorie für die Neuen Medien widersprechen.

I.

Bereits 1936 hat Walter Benjamin versucht, „Thesen über die Entwicklungstendenzen der Kunst unter den gegenwärtigen Produktionsbedingungen“ zu formulieren (zitiert nach der Ausgabe von 1963, Seite 10); er hat dabei Gedanken Bertold Brechts aus dem Jahr 1932 (wieder abgedruckt in der

Ausgabe Brecht 1992) aufgegriffen und weiterentwickelt. Die Untersuchungen von Benjamin und Brecht wurden in den späten sechziger und den siebziger Jahren weit rezipiert (vergleiche beispielsweise Hans Magnus Enzensberger 1970), spielen inzwischen im medientheoretischen Diskurs aber keine nennenswerte Rolle mehr. Dies ist unter anderem deswegen bemerkenswert, weil Benjamin und Brecht (oder auch Enzensberger) durchaus als scharfsinnige Autoren (noch immer) anerkannt sind und ihre medientheoretischen Überlegungen auch für die Gegenwart mehr als nur historisch interessante Erkenntnisse enthalten.

Offensichtlich liegt der Grund dafür, daß entsprechende Ansätze nicht mehr rezipiert werden, in scheinbaren oder tatsächlichen Fehleinschätzungen theoretischer wie politischer Art. Die Entwicklung schien anders verlaufen zu sein, als von den genannten Autoren vermutet und vorhergesagt worden war.

Hier soll untersucht werden, ob - und wenn ja: warum und inwieweit - dies tatsächlich der Fall ist - insbesondere vor dem Hintergrund auch der neueren Entwicklungen: der neuen Medien der Telekommunikation, insbesondere des Internet.

So kann als fraglich gelten, ob mit den scheinbaren oder tatsächlichen Fehleinschätzungen auch die politischen Vermutungen und Konsequenzen der genannten Autoren inzwischen (alle) obsolet geworden sind. Es gibt inzwischen mehr 'Apparaturen', um den von Bertold Brecht und Walter Benjamin benutzten Begriff aufzugreifen (derjenige des Mediums stand ihnen noch nicht zur Verfügung), als zu ihrer Zeit; und obwohl Hans Magnus Enzensberger 1970 bereits auf den Computer hinweisen konnte, liesen sich Entwicklungen wie die der Hypermedialität damals noch nicht abschätzen. Die Überprüfung der Thesen von Benjamin, Brecht und Enzensberger bezieht sich also auf mehr und andere Medien, die auch andere formale Charakteristika aufweisen, als die damals bekannten. Fraglich ist nun, in wieweit diese 'neuen Medien' mit den Thesen der genannten Autoren kompatibel beziehungsweise konsistent sind.

Zunächst wird deshalb die Position der genannten Autoren dargestellt. In einem zweiten Schritt wird untersucht, welche Auswirkungen die modernen Formen der Telekommunikation, insbesondere das Internet, vor dem Hintergrund der zuvor herausgearbeiteten Positionen aufweist. Schließlich führt diese Untersuchung zu eigenen Bewertungen im Hinblick auf eine gesellschaftspolitische Einordnung der neuen Massenmedien.

II.

Die Überlegungen Walter Benjamins, Bertold Brechts oder Hans Magnus Enzensbergers gehen auf die offensichtliche Erkenntnis zurück, daß die Medien Kunst und Unterhaltung, politische oder philosophische Texte, Meinungen und Darstellungen, ihre textliche, bildliche oder akustische Präsentation, selbst Gedanken und Gefühle allgemein verfügbar machen, wie auch

deren Träger und Protagonisten, teilweise unabhängig davon, was die inhaltliche Aussage ist. Benjamin schildert einen Demokratisierungs- und Egalisierungsprozeß, der den Medien - bereits der Schriftkultur, inzwischen beispielsweise auch dem Film - zu verdanken sei. Er beschreibt diese Chancen, die die (zu seiner Zeit: neuen) Medien bieten, mit poetischen Beispielen (1963. 15):

„Die Kathedrale verläßt ihren Platz, um in dem Studio eines Kunstfreundes Aufnahme zu finden; das Chorwerk, das in einem Saal oder unter freiem Himmel exekutiert wurde, läßt sich in einem Zimmer vernehmen.“

Theoretisch kann alles überall rezipiert und konsumiert werden - Ereignisse aus der Stadt auch auf dem Land, Weltkunst auf dem ehemals abgelegenen Dorf, politische Diskussionen aus den Zentralen und Hauptstädten auch in der Provinz; umgekehrt theoretisch auch Einschätzungen aus der Provinz in den Zentren. Daß dieser egalitären Hoffnung sozioökonomische und bildungsabhängige Wissenskluft-Effekte entgegenstehen, war zum Zeitpunkt, als auch noch Enzensberger seinen Artikel geschrieben hat, nicht bekannt (die Erstveröffentlichung der entsprechenden Erkenntnisse erfolgte ebenfalls 1970 durch Philipp J. Tichenor, George A. Donohue und Clarice N. Olien); sie widersprechen auch nicht der Tatsache, daß die entsprechenden Möglichkeiten jetzt erstmals in der Menschheitsgeschichte existieren, und sei es nur im Weberschen Sinne ‚idealtypisch‘ (Weber 1921) - wobei nicht auszuschließen ist, daß eine Verringerung der sozioökonomischen und bildungsabhängigen Unterschiede in der Bevölkerung auch eine verbesserte und egalitäre Rezeptionsbereitschaft entstehen lassen könnte.

Hans Magnus Enzensberger hält daher „die massenhafte Teilnahme an einem gesellschaftlichen [...] produktiven Prozeß“ (Enzensberger 1970. 160) für möglich. Dabei versteht er den Begriff der Teilnahme in einem weiteren Sinn und impliziert auch die aktive Partizipation am Mediengeschehen. Walter Benjamin äußert sich diesbezüglich gar euphorisch (1963. 32f.):

„So gibt zum Beispiel die Wochenschau jedem eine Chance, vom Passanten zum Filmstatisten aufzusteigen. [...]. Jeder heutige Mensch kann einen Anspruch vorbringen, gefilmt zu werden. [...]“

Dieses Diktum drückt eine politische Forderung, möglicherweise eine Utopie, vielleicht (aus damaliger Sicht) gar eine Erwartung, eine Prognose aus. Sie korrespondiert im übrigen auffällig mit Andy Warhols pointierter Formulierung von 1967: "In the future everybody will be world famous for fifteen minutes." Aber - im Gegensatz zur Warholschen Äußerung impliziert der Ansatz von Brecht, Benjamin oder Enzensberger konkrete politische Konsequenzen. Die Autoren gehen davon aus, daß die gesellschaftliche Teilnahme der ‚Massen‘, wie der damals gewählte Ausdruck lautete, über die Medien die Gesellschaft selbst verändere, demokratischer und egalitärer mache.

III.

Offensichtlich hat sich diese gesellschaftspolitische Erwartung nicht erfüllt; im übrigen scheint aber gerade dieser Sachverhalt es zu sein, die die Rezeption der Autoren heute erschwert.

Für die Fehlprognose der Autoren gibt es vermutlich mehrere Gründe; der wichtigste ist, daß die ‚Massen‘ auch über ihre ‚Teilnahme am gesellschaftlichen produktiven Prozeß‘, so er denn möglich ist, offenbar gesellschaftspolitische Inhalte nur bedingt geäußert oder auch nur konsumiert haben. Dies belegen zahlreiche Untersuchungen auch von Programmen, die unter Hinweise auf Brecht, Benjamin und Enzensberger konzipiert wurden und bei denen eine entsprechende Teilnahme nicht nur möglich, sondern gefordert ist (die sogenannten ‚Offenen Kanäle‘; vergleiche für eine solche Untersuchung Jarren/Grothe/Müller 1993).

Allerdings wäre es problematisch, aufgrund der politischen Fehleinschätzungen die Rezeption der Autoren insgesamt und die Diskussion ihrer theoretischen Vermutungen auf die Fehleinschätzungen zu verkürzen und einzustellen.

IV.

So muß beispielsweise und vor allem weiter diskutiert werden, ob eine ‚Teilnahme‘ im Sinn der Autoren grundsätzlich überhaupt möglich ist, und wenn ja, unter welchen Bedingungen und Einschränkungen.

Auffällig ist immerhin, daß selbst Walter Benjamin bei seinem Zitat (Benjamin 1963. 32f.)

„So gibt zum Beispiel die Wochenschau jedem eine Chance, vom Passanten zum Filmstatisten aufzusteigen. [...]. Jeder heutige Mensch kann einen Anspruch vorbringen, gefilmt zu werden. [...]“

noch von einer Gruppe ausgeht, die Filme herstellt; der ‚Passant‘ hat zunächst nur die Chance, zum Filmstatisten aufzusteigen. Zumindest sind professionelle Dritte notwendig, um ‚dem heutigen Menschen‘ sein Recht einzulösen, ‚gefilmt zu werden‘. Hier ist ein Spannungsverhältnis zu sehen, das aber von Benjamin, Brecht und Enzensberger wie auch in ihrer Folge eben nur unter einem politischen Gesichtspunkt thematisiert wird.

Tatsächlich muß gar eine Inkonsistenz festgestellt werden. Der Selbstverständlichkeit, mit der die Existenz professioneller Medienproduzenten akzeptiert wird, steht immer wieder die Vermutung entgegen, die Geschichte der Medien sei eine Geschichte zunehmender Teilnahme. So spricht Benjamin in seiner Analyse der schriftbezogenen Medien von dem langsamen Prozeß einer

wachsenden materiellen Verfügbarkeit und in der Konsequenz der Verbreitung und Ausdehnung entsprechender ‚Organe‘ (politischer, religiöser, wissenschaftlicher, beruflicher, lokaler Art), der auch zur Folge (gehabt) habe, daß

„immer größere Teile der Leserschaft - zunächst fallweise - unter die Schreibenden [gerieten]. Es begann damit, daß die Tagespresse ihnen ihren *Briefkasten* eröffnete, und es liegt heute so, daß es kaum einen im Arbeitsprozeß stehenden Europäer gibt, der nicht grundsätzlich irgendwo Gelegenheit zur Publikation einer Arbeitserfahrung, einer Beschwerde, einer Reportage oder dergleichen finden könnte.“ (Benjamin 1963. 33).

Fraglich ist aber, ob auch hier die abstrakte ‚Gelegenheit zur Publikation‘ mit dem tatsächlichen Zugang gleichzusetzen ist, oder ob ein solcher Zugang in der Praxis - und nicht nur aus Gründen des gesellschaftlichen Machterhalts - nicht zumindest bereits hier an weitere Bedingungen geknüpft wird. Diese Bedingungen spielen vor allem dann eine Rolle, wenn ein Distributionsprozeß eingeleitet und Rezipienten erreicht werden sollen.

Formalen, professionellen Kriterien muß aber offenbar jeder Genüge tun, der in einem ‚Organ‘ publiziert, das Leser gewinnen und halten will (wenn Wirkungen nicht erwünscht sind, wie beim Beispiel des privaten Tagebuchs, dann kann wiederum nicht von einem Kommunikationsprozeß gesprochen werden) - im übrigen auch ein ‚Organ‘, das gesellschaftliche Folgen anstrebt, nicht nur ein auf Gewinnmaximierung orientiertes Medium. So schreibt selbst Enzensberger (1970. 182), daß

„[i]n den osteuropäischen Ländern [...] die Nachrichtensprecher der Fernsehstationen viertelstundenlang Konferenz-Communiqués und ZK-Beschlüsse [verlesen], die nicht einmal für den Abdruck in einer Zeitung geeignet sind, offenbar in dem Wahn, sie könnten damit ein Millionen-Publikum fesseln.“

Die Mißachtung entsprechender ‚professioneller‘ Kriterien, hier: die Mißachtung einer zuschauerfreundlichen Mediendramaturgie, ist also nicht unbedingt nur bei sogenannten ‚Amateuren‘ zu beobachten; wenn auch diskutiert werden kann, ob die ‚professionellen‘ Medienschaffenden in Enzensbergers Beispiel gegen die Prinzipien einer zuschauerfreundlichen Mediendramaturgie nicht gerade deswegen verstoßen haben, weil Medienamateure ihnen entsprechende Anweisungen gegeben haben.

Entscheidend ist also, daß ein Medienprodukt hergestellt wird, welches Rezipienten finden muß. Partizipation ist in einem Massenkommunikationsprozeß - der Übermittlung von Inhalten von einer zentralen Stelle aus an viele Konsumenten - nur begrenzt möglich, allenfalls im Rahmen der von dieser zentralen Stelle (der Redaktion) vorgegebenen Standards, die sich wiederum an einem (möglicherweise in der konkreten Form dann nur angenommenen)

Geschmack des Zielpublikums, nicht einer Kleingruppe oder gar eines Individuums orientieren.

Solange es zentral hergestellte, zu vertreibende (distributierende) multipoint-to-point-„Organe“ gibt, ist es also aus formalen Gründen unrealistisch oder nur durch repressive Autoritätsausübung erreichbar, daß „[d]er Lesende [...] jederzeit bereit [ist], ein Schreibender zu werden [und] einen Zugang zur Autorenschaft [zu gewinnen]“ (Benjamin 1963. 33).

Insgesamt kann festgehalten werden, daß es auch aus anderen Gründen als denen der Bewahrung von Besitz- und Produktionsverhältnissen falsch ist, zu behaupten, die Unterscheidung zwischen Autor und Publikum sei im Begriff, ihren grundsätzlichen Charakter zu verlieren und zu einer funktionellen, von Fall zu Fall so oder anders verlaufenden zu werden (Benjamin 1963. 33).

Umgekehrt hat es bislang kaum Medien gegeben, die nicht diesen ‚Organ‘-Charakter aufgewiesen haben und die dezentral und auf individuell organisiert waren, aber dennoch distributiert werden konnten. Hans Magnus Enzensberger nennt lediglich den Kurzwellensender und kritisiert gerade an diesem Medium, daß es „in den Händen verstreuter radio hams zur harm- und folgenlosen ‚Freizeitgestaltung‘ heruntergekommen“ sei (Enzensberger 1970. 168). Tatsächlich weist der Kurzwellensender bereits zahlreiche formale Elemente der modernen Formen der Telekommunikation - wie insbesondere des Internet - auf; lediglich die Verbreitung ist deutlich schwächer.

Das Internet nun ist ein dezentral organisiertes Medium, das - so die These, die im folgenden erläutert werden soll - einem grundsätzlich demokratischen und egalitären Anspruch gerecht wird, ohne Selektions- und Redigierzwänge auskommt und trotzdem als Medium weit verbreitet und kommunikativ wirksam ist. Insbesondere soll hier die These vertreten werden, daß zumindest das Element der ‚Teilnahme‘ (allerdings ohne seine politischen Implikationen) mit dem Medium Internet erreicht werden kann.

Die damit unterstellte zweite - und in diesem Kontext weitaus gravierendere - Fehleinschätzung der genannten Autoren liegt demnach darin, ihre Erwartungen auf ein zentral organisiertes, publiziertes und distributiertes ‚Organ‘ konzentriert zu haben (point-to-multipoint), während sie dezentrale, auf individuelle oder Kleingruppenkommunikation ausgerichtete Medien (point-to-point) als zu wirkungsschwach weitgehend ignoriert hatten. Dagegen steht die Vermutung, daß gerade die Dezentralität die tatsächliche Möglichkeit der ‚Teilnahme‘ beim Internet begründet; während die Produktionszwänge zentraler Medien gerade wegen der dort notwendigen Massenwirksamkeit eine individuelle Teilnahme unmöglich machen. Allerdings war diese Fehleinschätzung offensichtlich erst dann zu erkennen, als dezentrale Medien eine derartige Verbreitung erreicht hatten, um über das Maß einer ‚harm- und folgenlosen Freizeitgestaltung verstreuter radio hams‘ hinauszuwachsen.

V.

Zumindest besteht für jeden Internet-Teilnehmer die *Möglichkeit*, sein eigenes Programmangebot zu gestalten und nicht nur zu ‚empfangen‘, sondern auch zu ‚senden‘.

Teilweise gibt es finanzielle Einschränkungen; und die Einschränkungen der Wissenskluft-Hypothese gelten auch und verstärkt für das Internet. Im folgenden konzentriert sich die Untersuchung also auf das Segment der Bevölkerung, das aus sozioökonomischen und bildungsbedingten Gründen Zugang zum Internet besitzt. Immerhin zeigen die Steigerungsraten der Internet-Nutzer, daß dieses Segment durchaus groß genug ist, um über das Maß einer möglicherweise vernachlässigbaren Gruppe wie die der 'radio hams' hinauszureichen.

Das Internet hat in einem weiteren Sinn Ähnlichkeiten zum Telefon, dem anderen großen dezentralen Medium für Individual- beziehungsweise (zunehmend) Kleingruppenkommunikation - tatsächlich dem einzigen point-to-point-Medium, das allgemein verbreitet ist. Wie beim Kurzwellenradio, gestatten auch Telefon und Internet die individuelle Wahl des Kommunikationspartners. Zudem gestattet und zwingt das Prinzip der Hyptertextualität, dessen Folge ein potentiell unendliches Verweisungsgefüge ergibt, dazu, das Internet mehr noch als das Telefonsystem und den Kurzwellenrundfunk als Meta-Kommunikations-Raum zu bewerten. Die Tatsache des Meta-Kommunikations-Raums ermöglicht dem Einzelnen die passive und aktive Teilnahme, ohne auf Massenwirksamkeit Rücksicht nehmen zu müssen. Jeder Teilnehmer kann sich einbringen, kommunizieren, mailen, in virtuellen Spielwelten Rollen übernehmen, sich darstellen: eben teilnehmen; offensichtlich wird die Möglichkeit zur Teilnahme auch intensiv genutzt und hat sich im Rahmen des Mediums durchgesetzt. Damit ist eins der Ziele erreicht, das von Autoren wie Brecht, Benjamin oder Enzensberger gefordert wurde. Paradoxerweise geschieht dies heute durch allgemeine Marktmechanismen, das damals ideologisch bekämpfte Prinzip, das nun zu einem Ergebnis führt, das ehemals gefordert worden ist, um den Markt zu verändern beziehungsweise zu beseitigen.

VI.

In wieweit treffen nun die gesellschaftlichen und politischen Vermutungen der Autoren auf diesen dezentralen Kommunikations-Raum zu?

Es ist auffällig, daß sich die Teilnahme überwiegend nicht auf politische Inhalte bezieht. Insofern wäre die Vermutung, politische Inhalte stünden als Folge der Teilnahme im Vordergrund, nachdrücklich zurückzuweisen. Damit entspricht das Verhalten der User im Internet demjenigen, das damals - um eine weitere

Dissonanz zu vermeiden - beklagt wurde, etwa von Hans Magnus Enzensberger (1970.170):

„Tonbandgeräte, Bild- und Schmalfilmkameras befinden sich heute schon in weitem Umfang im Besitz der Lohnabhängigen. Es ist zu fragen, warum diese Produktionsmittel nicht massenhaft an den Arbeitsplätzen, in den Schulen, in den Amtsstuben der Bürokratie, überhaupt in allen gesellschaftlichen Konfliktsituationen auftauchen. Indem sie aggressive Formen einer Öffentlichkeit herstellen, die ihre eigene wäre, könnten die Massen sich ihrer alltäglichen Erfahrungen versichern und aus ihnen wirksamere Lehren ziehen.“

Auch im Internet wird nur selten eine aggressive Form der Öffentlichkeit hergestellt. Die Mehrzahl der Beiträge bleibt im Bereich der Selbstdarstellung; auch der Freizeitbereich spielt eine wichtige Rolle. Offensichtlich konzentriert sich die Kommunikation und Teilnahme auf diese Themenfelder. Dabei gibt es spektakuläre Ausnahmen: so waren beispielsweise 1995 Protestlisten im Internet gegen die französischen Atomtests viel frequentiert, wobei wiederum einschränkend zu betonen ist, daß es sich hier um lediglich symbolische (einseitige) politische Aktionen gehandelt hat, bei denen eine Diskussion nicht notwendig geworden war. Zumindest gestattet das Internet also dem Individuum die Darstellung der eigenen Meinung; es ist daher, wie das Beispiel der Protestlisten gegen die französischen Atomtests gezeigt hat, ein geeignetes Medium für politische Lobbyarbeit. Alle Bereiche, in denen eine Meinungsbekundung in eine Richtung erfolgen kann und soll, können durch das Internet effizient und sinnvoll bedient werden. Fraglich ist jedoch, ob und in wieweit die politische und gesellschaftliche *Kommunikation* durch das Internet ermöglicht oder gefördert wird.

Die Dezentralität führt dazu, daß die Meinung eines jeden Teilnehmers gleich viel zählt - unabhängig von Herkunft oder formaler Bildung, auch unabhängig von der vertretenen Meinung oder dem Wahrheitsgehalt von Aussagen. Damit scheint das Internet ein demokratisches und egalitäres Medium zu sein. Wie aber wirken sich diese Prinzipien auf die politisch-gesellschaftliche Kommunikation aus?

Zunächst scheint sich ein quasi-demokratischer Abstimmungsprozeß zu entwickeln: Beispielsweise wird auf die WWW-Seiten, deren Darstellung (aus welchen Gründen im Detail auch immer) der herrschenden (öffentlichen, intersubjektiven) Meinung am nächsten kommt, am meisten zugegriffen, da auf sie mehr *links* führen, so daß auf sie öfter verwiesen wird. Je marginaler die Position ist, die auf den Seiten des Internet präsentiert wird, desto weniger *links* verweisen auf sie. Dabei bilden sich selbstreferentielle Kreise, die mit anderen Zirkeln kaum oder keinerlei Verbindung haben. Dieses Phänomen ist besonders auffällig in Foren: Im Umfeld eines Netzwerkes wird ein Vertreter einer oppositionären Meinung, und sei sie noch so gut begründet, häufig

pauschal diskreditiert oder schlicht ignoriert. So führt das Internet tatsächlich eher zu einem Nebeneinander nahezu tribalistischer Strukturen.

Die Dezentralität des Internet erschwert Formalismen, die eine Diskussion untereinander strukturieren würden, und die andererseits unabdingbar für politische Entscheidungsfindungsprozesse sind. Häufig ist das Nebeneinander so ausgeprägt, daß stereotype Meinungen noch weniger reflektiert, weil noch schneller und leichter reproduziert werden.

VII.

Das Internet ist also vor allem deswegen demokratisch, weil es egalitär ist und aus gleichwertigen Teilnehmern besteht, die nebeneinander existieren, deren Kontakte aber in der Regel folgenlos sind. Seine Form der Demokratie ist mithin nicht die eines kollektiven Miteinander, wie es den Positionen von Autoren wie Brecht, Benjamin oder Enzensberger entsprechen würde, sondern diejenige, die aufgrund dieses Nebeneinander von Individuen entsteht, das sich nicht vermeiden läßt und dem Medium immanent innewohnt. Der Begriff des Nebeneinander impliziert, daß Kommunikation nicht *notwendig* ist, sondern wenn, dann freiwillig mit anderen Individuen beziehungsweise Kleingruppen geschieht, mit denen nicht kommuniziert wird beziehungsweise werden muß, weil sie Nachbarn sind, sondern weil man sie sich erwählt hat - aufgrund gemeinsamer Interessen, undsoweiter. Es bedeutet auch, daß die Kommunikation wieder abgebrochen werden kann, ohne daß (negative) Folgen entstehen. Das Internet ist deshalb ein urbanes, individualistisches und individualisierendes Medium. Das Nebeneinander des Internet ermöglicht individuelle Freiräume und eine neue Form der Kommunikation im digitalen Raum.

Diese Form der Kommunikation und die durch sie gewährten Freiräume werden in der Regel als individuelle Chance und Bereicherung erlebt. Edgar Reitz schildert sie (1995. 275) in emphatischen Worten:

Ich bin innerlich mit [all den Menschen] verbunden. Wir sind das Netz. Und es ist international. Ich lebe nicht in München, sondern auf dem Planeten. Das ist es. Es entsteht ein planetarisches Gefühl [...].

Dieses metaphysische Gefühl *kann* aber nur entstehen, *weil* die Teilnehmer des Netzes keine persönlichen Kontakte, keine Alltagsprobleme miteinander haben, sondern die Kommunikation auf den freiwilligen Austausch über interessante Themen ‚reduziert‘ ist - weil ein persönlicher Kontakt nicht stattfindet.

Politik als lösungsorientiertes Handeln im gesellschaftlichen Kontext impliziert ein Ziel; um es zu erreichen, sind Rücksichtnahmen auf gesellschaftliche Zwänge notwendig, gegebenenfalls auch das Austragen von Konflikten. Wie

diese Rücksichtnahmen aussehen müssen, hängt von der gesellschaftlichen Situation ab, seien sie kultureller oder ökonomischer Art, sei es (lediglich) die aktuelle öffentliche Meinung, der Zeitgeist. In der Demokratie muß in der Regel die Form des Kompromisses akzeptiert werden. Es liegt auf der Hand, daß der Kompromiß nicht mit einem metaphysischen Gefühl, dem Eindruck eines universalen Eins-Seins korreliert.

Das Internet ermöglicht also den Freiraum der Individuation, es erweitert ihn. Diese Erweiterung erlaubt es dem Individuum, die Enge eines kollektiven Miteinanders hinter sich zu lassen. Dieses kollektive Miteinander aber bedingt Konflikte und Kompromisse. Es ist nachvollziehbar und entspricht der Theorie der kognitiven Dissonanz (Festinger 1962), daß Nutzer die Möglichkeit, Konflikte und Kompromisse vermeiden zu können, als Befriedung erleben und Strategien entwickeln (wollen), diese von subjektiven Impulsen bestimmte Form gesellschaftlicher Kommunikation beizubehalten und andere Formen zu vermeiden. Politik und Gesellschaft werden deshalb im Internet regelmäßig einseitig, 'bestenfalls' noch polarisiert erlebt, aber nicht als Notwendigkeit, Konflikte auszuhalten oder Kompromisse einzugehen.

VIII.

Zusammenfassend muß daher hinsichtlich der Argumentationskette von Brecht über Benjamin bis hin zu Enzensberger festgestellt werden, daß der Möglichkeit der Teilnahme an massenmedialer Kommunikation, von den genannten Autoren bereits für den Rundfunk, den Film undsoweiter postuliert, mit dem Medium des Internet tatsächlich keine (großen) materiellen oder technischen Widerstände entgegenstehen. Gerade die Dezentralität des Mediums ermöglicht die gesellschaftliche Partizipation. Das Internet ist tatsächlich ein egalitäres und weitgehend freies Medium.

Auch die zweite These der Autoren - mehr eine Erwartung: daß die Teilnahme an den Massenmedien zu politischen Änderungen führen könne und werde - scheint zuzutreffen, aber auch hier anders gerichtet, als von ihnen angenommen.

Hans Magnus Enzensberger war 1970 noch der Ansicht, daß die kollektive Nutzung der Medien zwangsläufig politisch sei und zu gesellschaftlichem Wandel führe. Andere Ziele bleiben „unpolitisch und borniert, sofern diese Produktion auf individuelle Bastellei hinausliefere“ (Enzensberger 1970. 168). Dieser Ansicht wird hier die These entgegengestellt, daß das Internet zwar egalitär und frei ist, in der Folge aber nicht notwendigerweise ein die Institution der Demokratie förderndes Medium darstellt.

Die Gegenposition soll wie folgt zusammengefaßt werden: Das Internet ist insoweit ambivalent, als es einerseits die Notwendigkeit zu kollektiven Gemeinschaften schwächt und zu einer individualistischen Gesellschaft führt -

dort aber wiederum neue Gemeinschaften ermöglicht, die nun auf Interesse, Freundschaft, immer aber subjektiven Faktoren gründet. Diese neuen Gemeinschaften werden intensiver; gleichzeitig schwächt sich die Bedeutung traditioneller kollektiver Gemeinschaften, die territorial begründet sind, immer mehr ab. Das Individuum erlebt diesen Prozeß als Befreiung - gerade auch von den Zwängen kollektiver Gemeinschaften, die Konflikte evozieren und Kompromisse erzwingen.

Dies bedeutet nicht, daß Politik kein Thema im Internet wäre oder sein könnte - im Gegenteil: es gibt Web-Pages und Foren, die ausschließlich politischen Themen gewidmet sind. Aber auch dort treffen sich überwiegend Gleichgesinnte, und es gibt keine gesellschaftliche Erfordernis, die ein kollektives Miteinander mit anderen erzwingen würde. Insgesamt bedeutet dies, daß das Internet *qua medium*, also nicht wegen seiner Inhalte, sondern aufgrund der formalen Bedingungen seiner Existenz für politische Kontexte nur begrenzt geeignet ist.

Fraglich muß an dieser Stelle bleiben, welche Auswirkungen dieser Sachverhalt für die Gesellschaft insgesamt hat - denn natürlich wird die Gesellschaft wie auch ihre politischen Strukturen nicht nur von der modernen Telekommunikation geprägt. Immerhin müßten die gesellschaftlichen Konsequenzen gerade der formalen Bedingungen neuer Medien besser untersucht werden. Daß gerade die formalen Aspekte des Mediums wichtiger als manche inhaltlichen sind, scheint immerhin deutlich geworden zu sein - und auch, daß die gesellschaftlichen Institutionen wie auch die politischen Strukturen in der Demokratie auf diese Entwicklung reagieren müssen.

Insofern entsprechen diese Überlegungen auch denen von Bertold Brecht, Walter Benjamin und Hans Magnus Enzensberger. Die Art der Reaktion gesellschaftlicher Institutionen und politischer Strukturen auf diese neuen medialen Zwänge muß aber notwendigerweise anders sein, als von den genannten Autoren erhofft beziehungsweise erwartet worden ist. Eine gesellschaftliche beziehungsweise politische Theorie der Telekommunikation oder des Internet muß erst noch formuliert werden.

References:

[Benjamin 1963]

Walter Benjamin, *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit*. In: Zeitschrift für Sozialforschung Vol. 1, 1936. Zitiert nach der Ausgabe: Walter Benjamin, *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit*. Drei Studien zur Kunstsoziologie. Frankfurt am Main 1963, S. 7 - 64

[Brecht 1992]

Bertold Brecht, *Der Rundfunk als Kommunikationsapparat*. 181ff. In: Blätter des Hessischen Landestheaters. Darmstadt: Nr. 16, Juli 1932, S. 181 - 184. Zitiert nach der Ausgabe: Bertold Brecht, *Schriften 1. Bertold Brecht, Werke*, Band 21. Frankfurt am Main; Berlin und Weimar: 1992, S. 552 - 557

[Enzensberger 1970]

Hans Magnus Enzensberger, *Baukasten zu einer Theorie der Medien*. In: Kursbuch 20 („Über ästhetische Fragen“), März 1970, S. 159 - 186

[Festinger 1962]

Leon Festinger, *A Theory of Cognitive Dissonance*. Stanford 1962. First published 1957

[Jarren/Grothe/Müller 1993]

Ottfried Jarren; T. Grothe; R. Müller, *Bürgermedium Offener Kanal*. Berlin 1993

[Reitz 1995]

Edgar Reitz, *Die Wiedergeburt des Epos aus den Netzwerken*. In: Edgar Reitz, *Bilder in Bewegung*. Reinbek bei Hamburg, November 1995, S. 222 - 267

[Tichenor/Donohue/Olien 1970]

Philipp. J. Tichenor; George A. Donohue; Clarice N. Olien, *Mass Media Flow and Differential Growth of Knowledge*. In: Public Opinion Quarterly Vol. 34; 1970, S. 159 - 170

[Warhol 1967]

Moderna Museet, Stockholm (Hrsg.), *Andy Warhol*. Stockholm 1967

[Weber 1921]

Max Weber, *Wirtschaft und Gesellschaft*. Tübingen: 2 Bände 1921

Optimale Informationsvorenthaltung als Strategem wissenschaftlicher Kommunikation

Gerhard Fröhlich

Institut für Philosophie und Wissenschaftstheorie
Johannes Kepler Universität Linz,
A-4040 Linz/Österreich
<gerhard.froehlich@iwp.uni-linz.ac.at>

Inhalt

- 1 Der 'homo academicus communicans' als fragwürdige Grundannahme
- 2 Die theoretisch-empirische Untersuchung wissenschaftlicher Kommunikation am Schnittpunkt von normativer Wissenschaftstheorie, Wissenschaftsforschung und Informationswissenschaft
- 3 Wissenschaftliche Institutionen - Garanten wissenschaftlicher Rationalität?
 - 3.1 Informationsblockaden in und zwischen Laboratorien
 - 3.2 Informationsvorenthaltung in wissenschaftlichen Publikationen
 - 3.3 Das kommunikative Elend der Kongresse oder: Rührende Unbeholfenheit als Strategem
- 4 Rationalere Wissenschaftskommunikation durch Digitalisierung?

Zusammenfassung

Funktionstüchtige Wissenschaftskommunikation, offene Kritik, freie kognitive Konkurrenz sind kein Beiwerk 'eigentlicher wissenschaftlicher Methoden', sondern für wissenschaftliche Rationalität unverzichtbar. Die weit verbreitete optimistische Grundannahme, daß Wissenschaftler andere Wissenschaftler bestmöglich informieren möchten und dies Sinn und Zweck der Wissenschaftskommunikation sei, sollte kritisch untersucht werden, v. a. auf der Basis von Konzepten und Befunden aus der Wissenschaftsforschung.

Abstract

Science and scientific objectivity do not (and cannot) result from attempts of an individual scientist to be 'objective', but from the „friendly-hostile“ (Popper) co-

operation of many scientists. The public character of scientific methods is constituted by free criticism and free cognitive competition. But the optimistic belief that scientists want to inform each other efficiently must be questioned, based upon concepts and findings of social studies of science.

1 Der 'homo academicus communicans' als fragwürdige Grundannahme

Strategem, lt. Duden „Kriegslist, Kunstgriff, Trick“; lt. The Oxford English Dictionary (1933, Vol. X, Oxford) u. a. „an artifice or trick designed to outwit or surprise the enemy; any artifice or trick; a device or scheme for obtaining an advantage; in generalized sense: skill in devising expedients; artifice, cunning“.

„Im 17. Jahrhundert, ja noch im 19., wurden Entdeckungen manchmal in Form eines Anagramms bekannt gemacht - so etwa Galileis 'Dreistern' Saturn und Hookes Elastizitätsgesetz -, um gleichzeitig die Priorität zu sichern und die Rivalen nicht auf Fährten zu locken, ehe der Gedanke weiter ausgebaut war. ... Es ist auch seit langem üblich, versiegelte und datierte Manuskripte bei wissenschaftlichen Akademien zu hinterlegen, um die Priorität wie auch den Gedanken zu schützen.“ (Robert K. Merton 1972, 145)

Sowohl in offiziellen Selbstdarstellungen von Wissenschaften als auch in Wissenschaftstheorie, Teilen der Wissenschaftsforschung und in der Informationswissenschaft ist folgende Grundannahme weit verbreitet - meist implizit, mitunter aber auch explizit formuliert: *WissenschaftlerInnen möchten ihre KollegInnen, die Mitglieder der sog. 'scientific community', bestmöglich informieren.* Wissenschaften seien kollektive Unternehmen, alle Forscher speisen demnach ihr Wissen unentwegt in den großen Wissensfundus der Menschheit ein, stellen sich schonungsloser Kritik und freier kognitiver Konkurrenz, sind so „Zwerge auf den Schultern von Riesen“ (Newton, Merton). Dies sei Sinn und Zweck der Wissenschaftskommunikation, v. a. der Kongresse und der wissenschaftlichen Zeitschriften. Dieser Grundannahme der ubiquitären Verbreitung der Norm des „Kommunismus“ in den Wissenschaften - neben dem Universalismus Kern des „wissenschaftlichen Ethos“ (Merton) - möchte ich im folgenden pointiert die These von der optimalen Informationsvorenthaltung als *Strategem* wissenschaftlicher Kommunikation entgegenstellen.

2 Die theoretisch-empirische Untersuchung wissenschaftlicher Kommunikation am Schnittpunkt von normativer Wissenschaftstheorie, Wissenschaftsforschung und Informationswissenschaft

Forschungs- bzw. Wissenschaftskommunikation¹ könnte als (potentielles) interdisziplinäres Forschungsgebiet am Schnittpunkt von Wissenschaftstheorie,

¹ Unter *Forschungskommunikation* fasse ich all jene kommunikativen Prozesse, welche von der ersten Idee bis zu den 'eigentlichen' wissenschaftlichen Produkten führen, zu neuen Befunden, Modellen, Theorien, mithin idealtypisch zum umfangreichen Forschungsbericht als Abschluß eines Projekts. Unter *Wissenschaftskommuni-*

Wissenschaftsforschung und Informationswissenschaft zu einer verstärkten Zusammenarbeit und gegenseitigen Befruchtung dieser Disziplinen führen:

(a) Daß für die *Informationswissenschaft*, die sich der Untersuchung und Optimierung von "Informationsprozessen in Wissenschaft, Gesellschaft, Wirtschaft und Verwaltung" (Wolf Rauch 1988, 3) widmet, Wissenschaftskommunikation ein zentraler Forschungsgegenstand ist, bedarf wohl keiner näheren Erläuterung.

b) In der *Wissenschaftsforschung*, der (Rück-)Anwendung der verschiedenen empirisch-theoretischen Forschungsmethoden auf die Wissenschaften selbst (als Wissenschaftssoziologie, -psychologie, -statistik etc.) finden sich zahlreiche disparate Modelle und Befunde zur Wissenschaftskommunikation. Nicht zuletzt aufgrund höchst unterschiedlichster disziplinärer Provenienz mangelt es bei ihnen vorrangig an systematischen Synopsen und Synthesen, an wissenschaftstheoretischer Fundierung wie theoretischer Integration.²

(c) Die normative *Wissenschaftstheorie* betrachtet Kommunikation und Information innerhalb der Wissenschaften keineswegs als nebensächliches Beiwerk zur "eigentlichen wissenschaftlichen Arbeit". Von allen ernstzunehmenden wissenschaftstheoretischen Konzeptionen der Gegenwart wird *Intersubjektivität*, wird der wissenschaftlichen Qualitätskontrolle über freie kognitive Konkurrenz und ungehemmte Kritik ein zentraler Stellenwert beigemessen.

So betont der Begründer des Kritischen Rationalismus, Sir Karl Popper (1969, 1970), die *sozialen, kommunikativen und institutionellen Aspekte* des Wissens und der Wissenschaften (vgl. Fröhlich 1998a,b):

(a) Popper stellt ausdrücklich der *Irrationalität der einzelnen Wissenschaftler* die *Rationalität der wissenschaftlichen Institutionen* gegenüber. Die Wissenschaft sei *keineswegs* auf der Unparteilichkeit oder Objektivität des individuellen Wissenschaftlers gegründet. Nicht nur Kultur- und Sozialwissenschaftler, gerade auch Naturwissenschaftler seien von einer „starken Parteilichkeit“ zugunsten ihrer intellektuellen Erzeugnisse“ beeinflusst. Die „wissenschaftliche Objektivität“ (diesen Ausdruck setzt Popper selbst ausdrücklich unter Anführungszeichen) sei daher nur als „Intersubjektivität der wissenschaftlichen Methode“ zu verstehen.

(b) Zwei Aspekte der "Methode der Naturwissenschaften" bilden nach Popper den *"öffentlichen Charakter der wissenschaftlichen Methode"*, nämlich *freie Kritik* und, so könnte man moderner formulieren, *gelungene Kommunikation*: "Zuerst gibt es so etwas wie *freie Kritik*. Ein Wissenschaftler mag seine Theorie mit der vollen Überzeugung ihrer Unangreifbarkeit vorbringen. Dies wird seine wis-

kation möchte ich alle kommunikativen Prozesse und Prozeduren fassen, die ab der Erstellung der 'Substanz' des Endberichts stattfinden, v. a. seine salamischeibenweise' Vermarktung im formalen wissenschaftlichen Publikationssystem, vgl. dazu meine Habilitationsschrift (Fröhlich 1998a). Forschungs- und Wissenschaftskommunikation werden in diesem Paper unter „wissenschaftliche Kommunikation“ zusammengefaßt.

² Quellenangaben können hier aus Platzgründen nur exemplarisch gemacht werden, vgl. dazu Fröhlich 1994/1998a sowie extensiv ebenfalls Fröhlich 1998a.

senschaftlichen Kollegen nicht unbedingt beeindrucken; es fordert sie vielmehr heraus. Denn sie wissen, daß die wissenschaftliche Einstellung darin besteht, daß man *alles kritisch untersucht*, und sie werden *selbst von Autoritäten nur wenig eingeschüchtert*. Zweitens *versuchen die Wissenschaftler, nicht aneinander vorbeizureden*. ... Selbst wenn sie verschiedene Muttersprachen verwenden, so versuchen sie doch sehr ernsthaft, eine und dieselbe Sprache zu sprechen." (Popper 1970, 267)

Den durch und durch sozialen und intersubjektiven Charakter der Wissenschaften, die nur *von Menschen im Plural* betrieben werden können, verdeutlicht Popper anhand der Situation von Robinson Crusoe (vor, so sei hinzugefügt, dem Eintreffen von Freitag). Selbst wenn es Robinson Crusoe gelungen wäre, auf seiner Insel Laboratorien oder astronomische Beobachtungsstellen einzurichten und auf Beobachtung und Experiment beruhende Abhandlungen zu verfassen, und selbst wenn Crusoes Konstruktion und Beschreibung von wissenschaftlichen Systemen mit den gegenwärtig von Wissenschaftlern akzeptierten Ergebnissen zusammenfallen würden, seien Crusoes Resultate keine Wissenschaft, sondern eher einer Offenbarung, einem Traum eines Hellsehers ähnlich: "Denn niemand außer ihm selbst korrigiert die Vorurteile, die die unvermeidliche Folge seiner besonderen geistigen Geschichte sind; niemand hilft ihm, jene seltsame Blindheit in bezug auf die inhärenten Möglichkeiten seiner eigenen Resultate loszuwerden, die sich einstellen muß, weil die Resultate zumeist auf relativ irrelevante Weise erhalten wurden ... nur beim Versuch, sein Werk *jemandem* zu erklären, *der es nicht ausgeführt hat*, kann er die Disziplin klarer und vernünftiger Kommunikation erlangen, die auch ein Teil der wissenschaftlichen Methode ist." (ebd., 269f)

Funktionstüchtige Wissenschaftskommunikation - v. a. in Form freier, "rücksichtsloser" (Popper) Kritik - ist mithin für Popper kein bloßer, nebensächlicher Zusatz zu einer 'eigentlichen' wissenschaftlichen Methode, sondern *zentraler und unverzichtbarer Bestandteil der wissenschaftlichen Methode*, wissenschaftlicher Rationalität. "Offenheit für Kritik" ist *das Kriterium für wissenschaftliche Rationalität*. Die wissenschaftliche Rationalität beruht mithin nicht auf der Rationalität von Einzelpersonen, sondern auf der Rationalität der wissenschaftlichen Institutionen, der wissenschaftlichen Kommunikation. Poppers Aussagen sollten allerdings nicht als deskriptiv, sondern müssen *normativ* verstanden werden: Die Funktionstüchtigkeit wissenschaftlicher Kommunikation kann nicht naiv vorausgesetzt werden, sondern die wissenschaftlichen Institutionen und Prozeduren sind laufend kritisch zu überprüfen, zu kritisieren und reformieren. Bestrebungen zu kognitiven und sozialen Protektionismen (Verhinderung von Kritik und freier kognitiver Konkurrenz) sind laufend und konsequent zu bekämpfen.

3 Wissenschaftliche Institutionen - Garanten wissenschaftlicher Rationalität?

Sir Popper hat eine institutionelle Theorie der Wissenschaftsentwicklung und des wissenschaftlichen Fortschritts entwickelt. Auf der Suche nach Bedingungen für den Fortschritt schlägt Popper ausdrücklich eine *institutionelle Analyse*

der Bedingungen des Fortschritts vor. Zu diesem Zweck regt Popper an, uns Bedingungen vorzustellen, unter denen der wissenschaftliche Fortschritt zum Stillstand käme - nämlich durch „Schließung oder strenge Überwachung“ von Forschungslaboratorien, wissenschaftlichen Zeitschriften und Kongressen, durch Schließung der Universitäten und Verbot von Büchern, Druckereien, Schreiben und Reden: „Alle diese Dinge, die tatsächlich unterdrückt (oder gelenkt) werden können, sind soziale Institutionen. Die Sprache ist eine soziale Institution, ebenso die Organisationen, die den Druck und die Veröffentlichung besorgen, und alle anderen Werkzeuge der wissenschaftlichen Methode. Die wissenschaftliche Methode selbst hat ihre sozialen Aspekte. Die Wissenschaft, und insbesondere der wissenschaftliche Fortschritt, ist nicht das Ergebnis isolierter Leistungen, sondern der *freien Konkurrenz des Denkens*. Denn die Wissenschaft braucht immer mehr Konkurrenz zwischen Hypothesen und immer rigorosere Prüfungen. Und die konkurrierenden Hypothesen müssen durch Personen vertreten werden: sie brauchen Anwälte, Geschworene und sogar ein Publikum. Diese persönliche Vertretung muß institutionell organisiert werden, wenn sie verlässlich funktionieren soll. Und diese Institutionen müssen unterhalten und gesetzlich geschützt werden.“ (Popper 1969, 120f.)

Untersuchen wir - in den engen Grenzen eines Kongreßpapiers unvermeidlich skizzen- und bruchstückhaft - etwas näher die Leistungsfähigkeit dreier wissenschaftlicher Institutionen, nämlich jener von Karl Popper explizit genannten "sozialen Institutionen, die zur Förderung der wissenschaftlichen Objektivität oder Unparteilichkeit ersonnen wurden; zum Beispiel der Laboratorien, der wissenschaftlichen Zeitschriften, der Kongresse." (Popper 1970, 268)

3.1 Informationsblockaden in und zwischen Laboratorien

Es fällt auf, daß Popper Laboratorien im Kontext der Kritik, der Kommunikation, des sozialen Charakters der wissenschaftlichen Methoden erwähnt. Dürfen innerhalb dieser und zwischen diesen Laboratorien funktionstüchtige kommunikative Verhältnisse unterstellt werden? Die Befunde der neueren qualitativen Wissenschaftsforschung, welche einen Schwerpunkt auf *Laborstudien* gelegt hat (vgl. z. B. Latour/Woolgar 1979, Lynch 1985, Knorr-Cetina 1984), sind äußerst ernüchternd - gerade den hier diskutierten Kontext rationaler Kommunikation und Kritik betreffend:

(a) Informationsvorenthaltung ist demnach sowohl *zwischen* Laboratorien als auch *innerhalb* von Laboratorien übliche Praxis: Nicht nur zwischen den Forschern verschiedener Laboratorien, sondern auch zwischen jenen Forschern (gerade auch im *Mittelfeld* der Statushierarchien), die an ein- und demselben chemischen Brüter oder Teilchenbeschleuniger arbeiten, werden eifersüchtig wichtige Ideen, entwickelte Methoden, Versuchsanordnungen, relevante neue Daten usw. versteckt und verheimlicht (vgl. Knorr-Cetina 1984).

(b) Auch *renommierte* Wissenschaftler, die zumeist in der Hoffnung engagiert werden, daß sie wertvolle Impulse, Motivierungen und Wissen an KollegInnen und Mitarbeiter weiterleiten, verfolgen zumeist ein *eigentümlich mehrgleisiges Informationsstrategem* (vgl. Beniger 1988), nämlich das *Innere* der Organisati-

on informationell abzuschöpfen, aber kaum akkumulierte (handlungs-)relevante Informationen ins Innere abzugeben, *nach außen* hingegen scheinbare Freigiebigkeit und Zuvorkommenheit zu demonstrieren. Dieses Grundmuster kann folgendermaßen beschrieben werden: *Informationsblockade nach innen* (gegenüber organisationsinternen Konkurrenten und statusniedrigeren Mitarbeiterinnen); *Absaugen der Informationen, Kompetenzen und sonstigen Handlungsressourcen aus dem Inneren der Organisation*; ihre (scheinbar, denn höchst selektive) teilweise 'altruistische' Verteilung per Post, E-Mail, Kongreßbeiträgen und Publikationen *nach außen*: Man mailt der Kollegin nach Australien, was man dem Kollegen im Nebenzimmer nicht gönnt. Der Ort größter Nachbarschaft (räumlicher oder inhaltlicher) ist der Ort größter Konkurrenz (Bourdieu) und daher auch oft der Ort größter Nicht-Kommunikation.

Wozu dient dieser Anschein von Kommunismus (Merton), diese mildtätig-großzügige Freigiebigkeit nach außen? Der *symbolische Mehrwert* der Wissenschaftskommunikation (vgl. Fröhlich 1994 bzw. 1996a), das 'symbolische Kapital' (Pierre Bourdieu) *wissenschaftlicher Reputation* ist die höchste Form wissenschaftlichen Sozialkapitals. Soziales Kapital, d. h. der Reichtum an Handlungsressourcen, die aus der Zugehörigkeit zu Beziehungsnetzen resultieren, und v. a. Symbolkapital (Anerkennung, Prestige), das aus einer hohen Statusposition in diesen Beziehungsnetzen resultiert, beruht nicht zuletzt auf einer *Geschenkökonomie*:

(a) So dienen auch im Wissenschaftsbetrieb (*Werbe*)-*Geschenke* diversester Art der Herstellung von Reziprozität, der Akkumulation von Verpflichtungskapital: Preprints, Gratisexemplare von Büchern, Daten, Einladungen zur Teilnahme an Großprojekten, wertvolle handlungsrelevante Informationen (z. B. Ausschreibungen/Terminfristen von Stellen, Forschungsförderungsschwerpunkten vor ihrer offiziellen Ausschreibung, ungeschminkte Situationsdiagnosen statt euphemistischer kommissioneller Protokolle), Einladungen zu Gastvorträgen bzw. Gastprofessuren, Verleihung von Ehrendoktoraten und Wissenschaftspreisen.

Diese Geschenke fungieren als eine Art *Wahlwerbung* - z. B. zur Förderung der eigenen Zitationsraten, welche den Ergebnissen der Stimmabgaben bei Wahlen in gewisser Hinsicht durchaus funktional äquivalent sind; im WWW entspricht dem die Vergabe von links. Zumindest muß ein großer Teil offiziöser Wissenschaftskommunikation als *P.-R.-Arbeit* angesehen werden, gerade auch die Informationsangebote von Institutionen und Einzelwissenschaftlern im World Wide Web. Die Logik dieser Geschenkverteilung ähnelt der Logik des demonstrativen altruistischen Luxuskonsums in vorindustriellen Gesellschaften (z. B.: pompöse, mehrtägige Hochzeitsfeiern). Warum? Auch das wissenschaftliche Schichtungssystem basiert auf *Ehre*, ebenso wie Rangpositionen in vorindustriellen Gesellschaften. Es dominiert daher eine ähnliche Logik der Praxis, die auf scheinbar irrational-freigiebig-altruistische Weise den eigenen Nutzen (oder, analog zur vormodernen Familienehre - Ehre und Nutzen der eigenen Paradigmengemeinschaft) verfolgt.

Scheinbar freigiebig sind die Informationsstrategien nach außen insofern, als die Regel gilt: *So wenig wie möglich* (wissenschaftlich handlungs-)relevante Informationen *zu veröffentlichen*, um Kollegen daran zu hindern, allzu schnell

die Versuche nachzubauen und zum Überholsprint anzusetzen; *nur so viel* an Informationen freizugeben, wie unbedingt nötig ist, um den *Erstanspruch* (*priority claim*) erfolgreich erheben (bzw. verteidigen) zu können. Pointiert-metaphorisch formuliert, geht es darum, möglichst geringwertige wissenschaftliche 'Glasperlen' herzugeben und dafür möglichst wertvolle wissenschaftliche Diamanten, Edelhölzer und sonstige Kostbarkeiten einzutauschen.

3.2 Informationsvorenthaltung in wissenschaftlichen Publikationen

Was verschweigen Wissenschaftler unter strategischen Gesichtspunkten (hier v. a.: Benachteiligung der Konkurrenten, aber auch: Verheimlichung von Problemen, Rückschlägen, Irrwegen) in ihren wissenschaftlichen Publikationen?

(a) Sie verschweigen z. B. *wichtige Details* ihrer Versuchsreihen (vgl. die prozeßorientierten Analysen unter 300 Biochemikern und Toxikologen eines kalifornischen Laboratoriums von Karin Knorr-Cetina 1984, v. a. 234ff.). Wie wichtig 'local knowledge' (etwa: das Wissen um die Eigentümlichkeiten eines bestimmten chemischen Brüters, die vor Ort entwickelten wissenschaftlichen 'Betriebsitten', die nicht die nach außen artikulierten methodischen Vorgangsweisen) in diesem Konnex ist, zeigt allein das Faktum, daß Wissenschaftler, welche ein bestimmtes wissenschaftliches Verfahren übernehmen wollen, dies meist in Form eines Praktikums in einer Institution, einem Labor erlernen, welches hier eine 'Vorreiterrolle' spielt.

(b) Auch Dritte-Welt-Länder machten die Erfahrung, daß der Erwerb der Patentrechte und das Studium der Patentschriften (durch eigene kompetente Fachleute) ohne Anwerbung erfahrener Techniker nur selten zum Erfolg führte, aufgrund des hohen Anteils an nicht dokumentierten handlungsrelevanten Informationen bzw. skills (praktischen Fähigkeiten). Gründe bzw. Ursachen dieser Nicht-Dokumentation liegen in Strategien der Informationsvorenthaltung, vielleicht auch etwas in Betriebsblindheit und - nach Karin Knorr-Cetina und auch meiner eigenen Ansicht nur in eher geringem Maß - in 'prinzipieller' Nicht-verbalisierbarkeit sog. 'impliziten' Wissens (Michael Polanyi).

(c) Generell gehen WissenschaftlerInnen mit Informationen, die den Entdeckungskontext (context of discovery) betreffen, d. h. die Strategien und Tricks ihrer Forschungs-, Kommunikations- und sonstigen Handlungsstrategien - äußerst sparsam um, ja versuchen sogar aufwendig die Spuren ihrer generativen Tätigkeit zu verwischen (Joerges).

(d) Auch die Entwicklung von *Geheim-* bzw. *Privatsprachen* kann als Form der Informationsvorenthaltung interpretiert werden, zwecks Behinderung von Kritik und Entlastung von umfassender Konkurrenz. Eine Konferenz der Herausgeber physikalischer Zeitschriften (April 1997, Chicago), dokumentiert in 'Science', erbrachte harsche Selbstkritik: der Großteil der Beiträge sei „*communication fog*“, d. h. völlig unverständlich: In jeder kleinsten physikalischen Subdisziplin werde - sachlich nicht notwendig und nicht gerechtfertigt - eine eigene Sprache entwickelt. In Zukunft müsse auf die *Lesbarkeit* von Beiträgen geach-

tet werden - z. B. sollten die Texte nicht mehr zu einem Drittel aus Akronymen bestehen.

Ich möchte alle diese scheinbar dysfunktionalen Verhaltensweisen das *Strategem der optimalen Informationsvorenthaltung* in der Forschungs- und Wissenschaftskommunikation nennen. Ihr Impetus: *Nur so viel wie nötig* halbwegs relevante Inhalte informell preisgeben, um Kooperationen aufrechterhalten zu können; quantitativ so viel möglich, aber in der relevanten Substanz *nur so viel wie unverzichtbar nötig* publizieren, um den Prioritätsanspruch wahren zu können; *so wenig wie möglich handlungsrelevante Informationen* informell weitergeben und vor allem publizieren, um zu verhindern, daß Konkurrenten daraus Wettbewerbsvorteile ziehen könnten. Wertvolle Informationen sind Objekte der Verknappung (zwecks Wertsteigerung), Tauschobjekte, Geschenke und werden nicht wahllos in die Wissenschaftsöffentlichkeit verstreut - oder gar im anonymen, potentiell eigentumsfeindlichen Anarchismus des Internet:

Die bereits erwähnten ethnographischen Feldstudien von Knorr-Cetina (1984) zeigen, daß Forschungspapiere bei genauer und unvoreingenommener Untersuchung nicht als das angesehen werden können, was sie gewöhnlich für sich beanspruchen, nämlich *Forschungsberichte* zu sein. Sie berichten nicht über die tatsächlich stattgefundene Forschung, ihre Probleme, Um- und Irrwege - und schon gar nicht, um ihre intersubjektive Überprüfbarkeit oder gar Wiederholbarkeit zu ermöglichen.

Das Gros der zum Verstehen eines (neuen) Verfahrens oder gar zur Replikation eines Experiments erforderlichen Informationen werden *nur über extensive zusätzliche informelle Informationsgewinnung* verfügbar. Hier sind Personen bzw. Gruppen, die selbst wertvolle Kapitalia (Handlungsressourcen) besitzen, also (langfristig) *Gegengeschäfte* bzw. *-geschenke* anbieten können, sicherlich im Vorteil gegenüber Außenseitern ohne solche Objekte des "einfachen Warentausches" (Marx). Als solche können fungieren (a) erfolgreich akkumuliertes vergegenständlichtes Kulturkapital (Geräte, Einrichtungen) und (b) Symbolkapital (Reputation von Personen, Institutionen, Journalen, Buchreihen). (c) *Soziales Kapital* im Sinne Bourdieus (Ressourcen aufgrund der Zugehörigkeit zu Beziehungsnetzen, die laufend gepflegt werden müssen, strong und weak ties im Sinne von Granovetter) ist hier eine höchst relevante Handlungsressource: "Die Wissenschaftler .. mobilisieren ein Netz von Verbindungen und Beziehungen, um ihre Informationen zu ergänzen. ... Wissenschaftler setzen bei dem, was sich nun *als Suche nach und Bemühen um relevante Information* herausstellt, als Ergänzung der formellen Kommunikation des Papiers informelle Kommunikation ein." (Knorr-Cetina 1984, 239; Herv. G. F.)

3.3 Das kommunikative Elend der Kongresse oder: Rührende Unbeholfenheit als Strategem

Auch die (wenigen) Studien zum Thema sowie die zahlreichen eher beiläufigen Bemerkungen über Kongresse und sonstige wissenschaftliche Treffen in der philosophischen und wissenschaftlichen Literatur und in Wissenschafts-Satiren (allen voran Arthur Koestler 1972) zeichnen ein wenig schmeichelhaftes Bild:

(a) Das ritualisierte Abspulen in höchster Geschwindigkeit von standardisierten, mit Fremdwörtern und Akronymen gespickten Referaten unter künstlichem Zeitdruck, (b) ihre in der Regel rhetorisch, didaktisch etc. äußerst unprofessionelle Gestaltung, bei (c) chronischer Überschreitung der Vortragszeit auf Kosten der vorgesehenen Diskussionszeit, (d) mit Overhead-Folien in selbst in den vordersten Reihen kaum lesbarer kleiner Schrift, mit einer Fülle von Information (Wörtern, Zahlen, nicht in der Legende aufgelösten Symbolen), die von den Zuhörern in der gezeigten Geschwindigkeit weder rezipiert noch kritisch bewertet werden können, sind immer wieder Gegenstand der Klage - hinterücks.

Selten wurde dabei die Überlegung angestellt, ob hinter diesen scheinbaren Unfähigkeiten und Schwächen, hinter dieser fast liebenswürdigen, ja rührenden Hilflosigkeit und Nicht-Professionalität (die wohl das *Durchdrungensein* von "Geist" symbolisieren soll, welcher über die Nichtigkeiten des banalen Alltags - hier: einigermaßen gelingender Informationsübertragung - erhaben ist), hinter dieser "ewigen Wiederkehr des gleichen" (Nietzsche) Dilettantismus' nicht System steckt, nämlich ein System durchaus effektiver Strategien der (präventiven) Verteidigung, der *Immunisierung vor Kritik*, des 'kalten Krieges' mit dem wissenschaftlichen Publikum, d. h. den anwesenden KonkurrentInnen und potentiellen KritikerInnen.

Meine These: *Alle diese Präsentations'fehler' und 'Ungeschicklichkeiten' gehorchen zwei strategischen Zielen: der Erschwerung, dem Schutz, der Immunisierung vor Kritik, und der optimalen Informationsvorenthaltung vor KonkurrentInnen:*

(a) Die meist unnötig überkomplizierten, kaum lesbaren Folien (ohne vorher schriftliche Unterlagen zu verteilen) sollen *einschüchtern*, die Rezipienten an der Funktionstüchtigkeit ihrer Augen und ihrer geistigen Aufnahmefähigkeit zweifeln lassen, sie informationell *überlasten* und stressen, ebenso wie (b) die meist sachlich nicht gerechtfertigte kaum vorhandene Verständlichkeit (aufgrund der Überladung mit prestigeträchtigen Ausdrücken), v. a. aufgrund des raschen Vorlesens von gedrechselten Schachtelsätzen. (c) Langatmige Einleitungen, Entschuldigungen und Erklärungen, was im folgenden Referat *alles nicht* behandelt werden würde, fressen nicht selten zwei Drittel der offiziellen Redezeit auf, so daß für die eigentliche originäre Botschaft der Referenten kaum mehr Zeit bleibt und diese nur bruchstückhaft und in noch höherer Geschwindigkeit vorgetragen wird.

So wird mehr oder minder erfolgreich erschwert, daß die Zuhörer den Inhalt des Referates verstehen und kritisieren können. Leider wagen es Kongreßorganisatoren bzw. Sektionsmoderatoren nur selten, hier energisch einzugreifen und die Redner an der Überschreitung der Vortragszeit (objektives Handlungsziel: Verhinderung längerer kritischer Diskussion, aber auch Raub an Vortrags- und Diskussionszeit der folgenden ReferentInnen) und am stakkatohaften (oder monotonen) Ablesen von Manuskripten zu hindern oder gar auf lesbaren Folien (oder zumindest auf das Verteilen von Kopien dieser Folien an das Publikum) zu bestehen.

Nicht nur hinter vorgehaltener Hand auf den Fluren der Kongreßzentren (und beim abendlichen Umtrunk), sondern auch in diversen Studien zum Thema ist daher nachzulesen, daß die Abhaltung und besonders die Rezeption von Vorträgen *nur als lästige Pflichtübung, als Höflichkeit gegenüber gewissen Kolleginnen*, aber als ansonsten wenig fruchtbringend erlebt wird. Man fährt auf Kongresse, um die *Pausen* und besonders die *Abende* informell lukrativ zu verbringen: hier wird *Schimpf- und Lobklatsch* ausgetauscht, unter der Hand mitgeteilt, "was gerade angesagt ist". Unter dem Siegel der Verschwiegenheit werden Tips über gewinnbringende Forschungsthemen, kommende Schwerpunktprogramme von Forschungsförderungsorganisationen, demnächst auszusprechende Stellen, Lehrstuhlvertretungen etc. ausgetauscht - d. h. über all das, was tatsächlich in der Wissenschaftspraxis handlungsrelevant ist. Jene, die erst - wenn überhaupt! - über Zeitschrifteninserate oder Mailinglisten davon erfahren, sind rettungslos im Nachteil: Eine, wenn nicht die wichtigste Form der Informationsvorenthaltung ist schließlich *Informationsverzögerung*. Da oft nur die Pausen als produktiv erlebt werden, finden sich, nebenbei bemerkt, in der Literatur auch ernst gemeinte Vorschläge, auf die störenden Vorträge dazwischen überhaupt zu verzichten.

Meine kritischen Äußerungen sollten als Plädoyer verstanden werden, in künftige Überlegungen zu rationaler und rationaler Kommunikation in den Wissenschaften auch die *verbalen* Kommunikationsformen mit einzubeziehen und sowohl kritisch zu untersuchen (empirisch aufwendiger als bisher) und zu modernisieren, d. h. den Entwicklungsstand wissenschaftlicher Treffen auf das in anderen gesellschaftlichen Feldern, etwa in der Wirtschaft, bereits länger übliche Niveau anzupassen, wobei der Nachholbedarf in den Sozial- und Geisteswissenschaften sicherlich größer ist als in anderen Disziplinen.

Daß Kongresse in absehbarer Zeit durch *elektronische* zw. *digitale* Diskussionsformen völlig ersetzt werden, ist nicht zu erwarten: denn nur wenige möchten wohl auf das informelle Zusammensein im Bier- oder Weinkeller, bei dem die eigentlich relevanten Gespräche stattfinden, und auf den *Belohnungscharakter* dieses Kongreßtourismus' verzichten. Dem entspricht auch, daß nach einer Reihe von Untersuchungen über Internet-Kommunikation die *face-to-face* (etwa auf Kongressen oder bei Laborbesuchen) *begonnenen* und dann digital gestützt fortgesetzten *Beziehungen* die intensiveren und fruchtbareren wissenschaftlichen 'Fernbeziehungen' sein sollen. Kaum ein Wissenschaftler (oder gar ein Manager in der Wirtschaft) verrät Unbekannten, mit denen er noch nie einen Händedruck gewechselt (und, je nach Habitus, gepflegtes Bier, erlesenen Rotwein oder noblen Champagner getrunken) hat, wertvolle Ideen oder Verfahrenstricks. Die Botschaften der diversen offenen (unbeschränkt zugänglichen) Internet-Mailing-Listen sind daher auch in der Regel Versuche, die kumulierte Intelligenz und das aggregierte Gedächtnis der Listenteilnehmer an- und abzapfen, aber meist auf einem eher harmlos-philologischen Niveau (nach dem Strickmuster von "Wer kennt eine deutsche Übersetzung von..."). Das Klientel dieser Gruppe besteht primär aus Studentinnen und NachwuchswissenschaftlerInnen.

Mit anderen Worten zusammengefaßt: In Forschungs- und Wissenschaftskommunikation ist der Anteil der 'informellen' Kommunikation erheblich: Diese ist nicht einfach 'gratis' zu haben, sondern verlangt zumindest mittelfristig kom-

pensatorische Gegenleistungen; sie findet also insofern nicht 'ohne' Formen, sondern in *subtileren* Formen statt. Die offizielle Verlautbarung mittels Vortrag oder formaler Publikation ist sohin *nur als Spitze des Eisbergs relevanter Information* anzusehen. WissenschaftlerInnen ohne Sozialkapital, die aus relevanten informellen Cliquen, Seilschaften und Beziehungsnetzen ausgeschlossen sind, sind schwer benachteiligt.

Anonymität und Arkanpraxis des *Referee-Systems*, wohl eine der schärfsten Varianten der Informationsvorenthaltung, der asymmetrischen Beziehungen in der Wissenschaftskommunikation, wird in der Literatur häufig und heftig attackiert. So wissen WissenschaftlerInnen in der Regel weder, *wer* sie verurteilt (ist es etwa jener Kollege, den man letztes Jahr in einem Artikel vorsichtig zu kritisieren wagte und der jetzt vielleicht Gelegenheit zur Rache nimmt?), noch bekommen sie die über sie bzw. ihre Werke erstellten *Gutachten* zu Gesicht (bestenfalls erfolgt dies auszugsweise). Gutachter können sich anonym vom Ideenfundus der Einreichenden anregen lassen und diese später, selbst bei besten Vorsätzen, als eigene Ideen verkünden (sog. Kryptamnesien). Vielfach sehen die kritischen Studien zur Untersuchung der Leistungsfähigkeit des Peer-Review-Systems durch die derzeit üblichen Geheimpraktiken die Offenheit der Wissenschaftskommunikation bedroht.

4 Rationalere Wissenschaftskommunikation durch Digitalisierung?

Birgt die - schon allein ökonomisch unvermeidbare und begrüßenswerte, vgl. Fröhlich 1998a - konsequente Digitalisierung der Wissenschaftspraxis Potentiale für eine offenerere, kritischere und sohin *rationalere* Wissenschaftspraxis?

Eine rationalere Wissenschaftspraxis wäre eine, welche (a) offene und freie Konkurrenzverhältnisse (wieder) herstellt, weiter ausbaut und erfolgreich schützt, (b) freie „rücksichtslose“ (Popper) Kritik ermöglicht, also (c) kognitiven Protektionismus und den ihn stützenden sozialen Protektionismus einzudämmen versucht. Sie sollte also *für mehr kognitive und soziale Chancengleichheit im "Krieg der Ideen"* (Popper) sorgen. (d) Dieser Krieg sollte kein 'kalter Krieg' der Ideen (d. h. immer auch: der dahinter stehenden Individuen, Gruppen, Institutionen) bleiben - gerade auch mittels diversester unterschwelliger Manipulationen, Intrigen und Tricks betrieben -, sondern als offene, faire "freundlich-feindliche" (Popper) Auseinandersetzung geführt werden. Auch (e) leistungsfeindliche Effekte aufgrund der Konzentration und Zentralisation sozialen und symbolischen Kapitals (des sog. „*Matthäus-Effekts*“ nach Merton, vgl. Fröhlich 1994) sollten gemindert oder kompensiert werden. Können digitale Informationstechnologien im Sinne offener Kritik, des Theorienpluralismus, allgemein: der institutionellen Förderung und Schutzes freier wissenschaftlicher kognitiver Konkurrenz förderlich sein?

In der neueren Literatur (und in den Internet-Gruppen bzw. -Publikationen) zum Thema Informationsgesellschaft, zu den Effekten und Potentialen digitaler Informationstechnologien und zur Digitalisierung der Wissenschaftskommunikation wurden wiederholt übergroße sozialutopische Hoffnungen artikuliert. Sie lassen sich in der Hoffnung auf eine *umfassende Demokratisierung* aller Sekto-

ren, welche durch die digitalen Informationstechnologien tieferschürfend beeinflußt werden, zusammenfassen.³ Viele Autoren verbinden vor allem mit dem 'anarchistischen' Internet Hoffnungen hinsichtlich einer „Demokratisierung“ der Wissenschaftskommunikation. Sicherlich erschweren dezentrale Computernetze, Von-Vielen-an-Viele-Kommunikationsmodelle die Bildung von Informationsmonopolen.⁴ Auch die Artikulation von Kritik ist über die vielen themenspezifischen wissenschaftlichen Mailing-Listen leichter möglich. Es ist allerdings fraglich, inwiefern solche Kritik zur Kenntnis genommen wird, da sie zumindest vorderhand (noch?) als statusniedrig, als minderwertige informelle Kommunikation abseits des 'eigentlichen', formalen Publikationssystems (mit Peer-Review-Prozeduren) abgewertet wird.

Allgemein wird über die Verbreitung computerunterstützter Wissenschaftspraxis und Wissenschaftskommunikation die *Verringerung der Chancenungleichheiten im Zugang zu wissenschaftlichen Handlungsressourcen* erhofft - nicht zuletzt durch verbilligte und erleichterte, *rationellere* Wissenschaftspraxis. Rationellere Wissenschaftspraxis hieße Verminderung kollektiver wie individueller Kosten der Wissenschaftspraxis, u. a. über (a) Verhinderung unproduktiver Mehrfachforschung, Einsparung vermeidbarer Kosten, Automatisierung repetitiver Vorgänge (wie des ständig wiederholten Abtippens von Zitaten, Tabellen, Literaturangaben), direkte Eingabe von Daten bei Befragungen bzw. Befragung per Internet, (b) beschleunigten und erleichterten Zugang zu wissenschaftlichen Primärinformationen (= den gewünschten Originalarbeiten, die benötigten Daten, den direkten Kontakt zu anderen Forschern auf demselben Gebiet), (c) geringerer zumindest zeitlicher Arbeitsaufwand bei der Auswertung von Datenbank- bzw. Internet-Search-Engines-Recherchen, aber auch der akkumulierten Originalarbeiten mittels computerunterstützter Verfahren (z. B. mittels halbautomatischer Übersetzung, Einsatz von Retrieval- und inhaltsanalytischen Programmen, statistischer Auswertungen und Rangreihenbildungen von Datenbankrecherchen zwecks Übersichtsgewinnung), sowie (d) beschleunigtere und potentiell transparentere Peer-Review-Prozeduren mittels Umstellung der AutorInnen-Verleger-Gutachter-Interaktionen auf E-Mail-Basis und Preprint-Server.

Die Verbilligung und daher massenhafte Zugänglichkeit von wissenschaftlichen Produktions- und Kommunikationstechnologien ist sicherlich potentiell demokratisierend, erleichtert insofern Kritik und freien kognitiven Wettbewerb:

(a) Die frühere große optische Diskrepanz zwischen gedruckten, mit vielen Tippfehlern und erkennbaren Korrekturen 'befleckten' Typoskripten und professionell gesetzten Büchern ist fast völlig verschwunden, im Gegenteil: im WWW publizierte Beiträge übertreffen bereits optisch konventionell gedruckte Beiträge, v. a. aufgrund ihrer Multi-Media-Funktionen. Nun mag eingewendet werden, daß der optische Mehrwert einer Publikation keinerlei Zusammenhang zu ihrem kognitiven Wert habe und daher auch nicht in die Bewertung einfließen sollte. Dies ist selbstverständlich normativ zu unterstützen, gleichwohl dürfen die Effekte der "Verpackung" wissenschaftlicher Arbeiten nicht unterschätzt werden.

³ Vgl. zu eigenen Arbeiten, welche diese Hoffnungen darstellen und einer Kritik unterziehen, Fröhlich 1994/1996a, 1995, 1996b, 1997.

⁴ So wurde über das Internet eine geheime Studie der Tabakindustrie verbreitet, welche eindeutig den Suchtcharakter des Zigarettenkonsums belegte, sowie Strategien, die Sucht der Raucher durch künstlich erhöhte Nikotinbeigaben zu fördern.

Da bekanntlich ein Großteil der wissenschaftlichen Buchproduktion, aber auch viele wissenschaftliche Journale ökonomisch auf nicht unerheblichen Kostenbeiträgen der Autoren beruhen, ist eine (fast) kostenfreie Publikation im Internet ebenfalls ein Schritt zur Chancengleichheit.

(b) Bei geschickter Nutzung sinken die Kommunikations- und Organisationskosten über E-Mail und WWW-Server (im Vergleich zu konventioneller Vorbereitung mittels Fotokopien, Briefpost, Telefon und Fax) beträchtlich. Dies kann zweifellos auch finanzschwächeren Wissenschaftlern, Gruppen, Instituten die Organisation von Kongressen erleichtern. Auch ist es, zumindest auf der Basis zeitlich ausgedehnter ehrenamtlicher Aktivitäten, prinzipiell jedem möglich, kleinere Datenbanken und Informationsdienste auf verschiedenen Gebieten aufzubauen und anzubieten. Generell können auf der Grundlage des Internets und anderer Computernetze kostengünstig und vergleichsweise problemlos Forschungsvorhaben angegangen werden, welche geographisch distante Kooperation (z. B.: Meteorologie, Ozeanographie, Astronomie) erfordern oder aufgrund des hohen zeitlichen Aufwands eine arbeitsteilige Zusammenarbeit mit Konkurrenten erzwingen (Analyse des menschlichen Genoms). Solche Kooperationsprojekte werden allerdings - aus nachvollziehbaren Gründen - meist nicht öffentlich, sondern über geschlossene Listen abgewickelt. Zumindest bleibt der Zugang zum Datenmaterial während einer Sperrfrist nur den unmittelbaren Projektteilnehmern oder den Mitgliedern der Trägerorganisation (etwa einer astronomischen Gesellschaft) vorbehalten.

Der anarchistische „*Freie Fluß der Information*“, egalitär, öffentlich und kostenfrei, so wie er in der Internet-Literatur und in Internet-Diskussionsgruppen als konstitutiv für das Internet behauptet wird, bleibt weiterhin eine *harmonistisch-utopische* Verheißung. Auf dem Hintergrund sozialwissenschaftlicher Theorien, etwa der Feld- und Distinktionstheorie Pierre Bourdieus, ist es wesentlich realitätsgerechter, Computernetze als soziale Felder wie andere, mithin als *vertikal geschichtete Konkurrenzfelder* zu sehen: Auch in Computernetzen verfolgen Menschen Ziele, möchten Ansehen erringen, Gruppen bilden und andere von diesen ausschließen. So erfolgt, wie bereits vermerkt, im WWW über das Einrichten von links die Zuteilung von sozialem bzw. symbolischem Kapital, können sich Seilschaften (ähnlich den Zitationskartellen in der Wissenschaftskommunikation) gegenseitig fördern und andere benachteiligen, resp. ausgrenzen. Auch der raue Ton in vielen Diskussionslisten läßt nicht gerade auf sanft-ideale Kommunikationsgemeinschaften schließen.

Macht ist eine Struktureigentümlichkeit aller sozialen Beziehungen (Norbert Elias). Die Quelle von Macht ist die *Kontrolle über (Handlungs-)Ressourcen, die andere benötigen*. In den verschiedenen sozialen Konkurrenzfeldern ist daher weiterhin Informationsvorenthaltung, -blockierung, -verzögerung eine effektive Strategie. Handlungsrelevante (d.h. immer auch: zum richtigen Zeitpunkt vorliegende) Informationen, sowohl strukturelle als auch praktische Detailinformationen, sind weiterhin wertvolle, knapp gehaltene Güter. Es gilt für handlungsrelevante Informationen, daß ihre allgemeine Verbreitung ihren Handlungs- und Distinktionswert sinken (d. h. ihren Banalitätswert steigen) läßt. Informationen werden daher künstlich knapp gehalten bzw. als Geschenk und Belohnung eingesetzt, als teure Beratungsleistung verkauft. Professionen und andere mächtige Gruppen unserer Gesellschaft leben geradezu von der Zurückhaltung, der Monopolisierung von Information. Es ist mehr als fraglich,

warum - wie viele euphorische Berichte bzw. Zukunftsszenarios zum sog. 'Cyberspace' unterstellen - im digitalen Raum ein neuer Altruismus, ein Kommunismus des Wissens walten sollte.

Der freie, egalitäre, kostenfreie Fluß der Information, der freie Austausch der Ideen, hemmungslose Kritik und kognitiver Wettbewerb, ist mithin eine (realitätsferne) Utopie, bestenfalls eine regulative Idee. Es könnte jedoch eine fruchtbare gemeinsame Aufgabe von normativer Wissenschaftstheorie, quantitativer wie qualitativer Wissenschaftsforschung und der Informationswissenschaft sein, die *sozialen, institutionellen, informationellen Bedingungen wissenschaftlicher Rationalität* zu untersuchen und praxisrelevante Korrekturvorschläge zu erstellen. Ethische Appelle an die EinzelwissenschaftlerInnen werden nur mäßige Erfolge zeitigen, so lange die *Belohnungsmechanismen* der Wissenschaften offene Kritik und kognitive Konkurrenz, die Offenlegung von Fehlern und Problemen eher bestrafen als belohnen.

Auf die Eigendynamik digitaler Technologien allein können wir nicht vertrauen. Durchgreifende *antiprotektionistische wissenschaftspolitische Reformen* tangieren viele 'vested interests' und könnten daher nur gegen erheblichen Widerstand durchgesetzt werden. Sie wären als flankierende Maßnahmen unverzichtbar, um den Listen und Tücken optimaler Informationsvorenthaltung produktivere Handlungsimperative entgegenzusetzen, welche der Rationalität des wissenschaftlichen 'Gesamtarbeiters' förderlich sind. Da die vielfältigen Strategien der Informationsvorenthaltung als nicht antizipierte, unerwünschte Folgen der 'Bedrohung' der WissenschaftlerInnen durch Konkurrenz und Kritik anzusehen sind, müßten wir auch ein *fehlerfreundlicheres* (Popper) kulturelles Klima schaffen. Letztlich gibt es nämlich nur einen Weg der Streßvermeidung: sich und anderen unsere Fehler, Irrtümer und Mißerfolge zu verzeihen. So könnten Potentiale digitaler Technologien zur Förderung rationaler Wissenschaftskommunikation tatsächlich umfassend und wirksam genutzt werden.

Literatur

[Beniger 1988]

Beniger, J. R. (1988): *Information Society and Global Science*. The Annals of the American Academy of Political and Social Science 495, S. 14-28

[Fröhlich 1994]

Fröhlich, G., *Der (Mehr-)Wert der Wissenschaftskommunikation*, in: Rauch, W. (Hg.): *Der Mehrwert der Information*. Konstanz 1994, S. 84-95

[Fröhlich 1995]

Fröhlich, G., *Demokratisierung durch Datenbanken und Computernetze?* in: Becker, T. et al. (Hg.): *Informationsspezialisten zwischen Technik und gesellschaftlicher Verantwortung*. Stuttgart 1995, S. 55-60 bzw. <<http://www.uni-stuttgart.de/UNLuser/hbi/publikat/hbipubl/guides/froehlic.htm>>

[Fröhlich 1996a]

Fröhlich, G., *Surplus Values of Scientific Communication*. Review of Information Science, Volume I (Number 2) 1996, <<http://www.inf-wiss.uni-konstanz.de/RIS/>>

- [Fröhlich 1996b]
Fröhlich, G., *Netz-Euphorien. Zur Kritik digitaler und sozialer Netz(werk-)metaphern*, in: Schramm, A. (Hg.): *Philosophie in Österreich 1996*, Wien 1996, S. 292-306
- [Fröhlich 1997]
Fröhlich, G., *Mythos Informationsgesellschaft? Informationsblockaden und Desinformation in Wissenschaft und Gesellschaft*, in: Meleghy, T. et al., Hg., *Soziologie im Konzert der Wissenschaften*. Opladen 1997, S. 328-336
- [Fröhlich 1998a]
Fröhlich, G., *Kontrolle durch Konkurrenz und Kritik? Der öffentliche und soziale Charakter der wissenschaftlichen Methoden*. Linz 1998: Institut für Philosophie und Wissenschaftstheorie
- [Fröhlich 1998b]
Fröhlich, G., *Kontrolle durch Konkurrenz und Kritik? Der öffentliche und soziale Charakter der wissenschaftlichen Methoden*, in: Winfried Löffler, L./Runggaldier, E. (Hg.): *Vielfalt und Konvergenz der Philosophie*. Wien 1998
- [Koestler 1972]
Koestler, A., *Die Herren Call-Girls*. München/Zürich 1972
- [Latour/Woolgar 1986]
Latour, B./Woolgar, S., *Laboratory Life. The Social Construction of Scientific Facts*. Princeton, N.J. 1986
- [Lynch 1985]
Lynch, M., *Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory*. London etc. 1985
- [Merton 1972]
Merton, R. K., *Die Priorität bei wissenschaftlichen Entdeckungen*, in: Weingart, P., Hg., *Wissenschaftssoziologie I. Wissenschaftliche Entwicklung als sozialer Prozeß*, Ffm. 1972, S. 121-164
- [Knorr-Cetina 1984]
Knorr-Cetina, K., *Die Fabrikation von Erkenntnis. Zur Anthropologie der Naturwissenschaft*. Ffm. 1984
- [Popper 1969]
Popper, K., *Das Elend des Historizismus*. Tübingen 1969
- [Popper 1970]
Popper, K., 1970, *Die Wissenssoziologie*, in: ders.: *Die offene Gesellschaft und ihre Feinde*, Bd. II. Bern/München (c1958), S. 260-274
- [Rauch 1988]
Rauch, W., *Was ist Informationswissenschaft*, *Fakten Daten Zitate* 3/1988, S. 1-6

Autorenindex

Assfalg	385	Merschjan	80
Bauer	137	Milchrahm	88
Bayraktar	427	Nake	455
Bekavac	438	Odenthal	283
Borovansky	256	Panyr	235
Buckner	67	Rautenstrauch	168
Burke	305	Reichmann	137, 508
Cejpek	481	Rittberger	438
Currás	17	Rosický	469
Davenport	67	Schaefer	413
Dietzsch	201	Schneider	80, 404
Engel	273	Schweibenz	185
Esswein	201	Schweitzer	80, 404
Fröhlich	535	Semar	263
Giessen	523	Shafaei	32
Grossmann	105	Siegle	168
Hammwöhner	47	Siemonsen	273
Herget	105	Sigel	296
Heuer	273	Simsova	123
Hoechst	80	Skenderija	253
Hüther	397	Sosna	258
Kaiser	329	Staudinger	358
Kalkus	501	Steinerová	143
Kimlicka	213	Stempfhuber	413
Knorz	154	von Teutul	404
Krause	413	Vlasak	501
Krickl	88	Vodáček	58
Kuhlen	283	Wilson	225
Macevičūtė	486	Wolff	373
Mandi	312, 413	Womser-Hacker	427
Marx	358	Yetim	344
Meinel	273	Zucker	235